This lecture examines the case for intervention in consumer financial markets. The tension between laissez-faire and interventionist tendencies is as old as the discipline of economics itself. Laissez-faire economists appreciate and defend the performance of free markets, while interventionists identify market failures and argue that feasible policies can be found to correct them. The problems of greatest concern to interventionists vary over time, but famously include monopoly power since the late nineteenth century, aggregate demand management since the 1930s, and, over the last 100 years, consumer protection in various spheres including food and drugs, autos and consumer durables, and financial services.

Interventionists are often motivated by the desire to forestall more extreme “revolutionary” responses to social problems. A standard goal of interventionism is to design policies that restore the welfare properties that free markets would have if market failures were eliminated, thereby restoring the validity of classical economics. Thus, the trust-busters of the Progressive era sought to restore competition, Keynesian economists seek monetary and fiscal policies that stabilize the macroeconomy and restore textbook microeconomics—the “neoclassical synthesis” of Samuelson (1955)—and consumer regulators seek to restore the choices that consumers would make if they were rational and well informed.

Richard T. Ely, one of the founders of the American Economic Association, was a leading interventionist of his era. His chief concern was the problem of monopoly, but he also wrote about the unethical nature and destructive effects of competition in markets with poorly informed consumers.1 He argued that such

---

1 For example, Ely wrote in his *Introduction to Political Economy* (1894): “It is of no avail to say that business is excluded from the domination of ethical principles, for it is precisely in our economic life that ethical principles of any real validity must manifest themselves. It is only in an imperfect condition of society that sharp practice and hard bargaining can ever appear to men to be morally right”
problems needed to be addressed in order to preserve political support for a broadly free-market economic system.\footnote{A few pages later he writes that “Competition tends to force the level of economic life down to the moral standard of the worst men who can sustain themselves in the business community” (p. 83).}

Reading his work today, one is impressed by his broad interests and the contemporary nature of many of his concerns, but also by the lack of evidence he provides for the effectiveness of his proposals, some of which reflect prejudices of his day.\footnote{In The Strength and Weakness of Socialism, Ely writes: “If we allow things to take their own course, if we remain passive in the presence of the evils which socialism has so amply demonstrated and vividly depicted, the result may well be that outcome which the evolutionary socialism of Marx has pointed out. But there is no reason why we should remain passive in the presence of evils. On the contrary, there is every reason why we should vigorously attack existing evils, and so do with the hope that they can be abated and improvements in social conditions can be effected” (1899, p. 255).}

His life and work are an inspiration, but also a warning that we need to rely on strong evidence rather than intuition when proposing any intervention in the economy.

The subject of this lecture is household finance, also known as consumer finance. The financial decisions of consumers have attracted increasing interest in recent years as an area where competitive markets may deliver standard outcomes that can be improved through intervention. Within the last ten years, the reach of regulation has been extended by the most important US consumer financial legislation since the New Deal of the 1930s, notably the Pension Protection Act of 2006, the Credit Card Accountability Responsibility and Disclosure Act (CARD Act) of 2009, and the Dodd-Frank Act of 2010 that created the Consumer Financial Protection Bureau. Similar trends are visible around the world.

Household finance is a contemporary focus of interventionist attention for several reasons. Modern economies have evolved in a way that requires individuals to make more difficult financial decisions with bigger consequences. Most obviously, people are living longer, while traditional defined-benefit retirement systems that provide guaranteed income streams are being replaced with defined-contribution systems that require people to accumulate and invest their own retirement savings. Higher education is becoming increasingly expensive everywhere, posing a financing challenge in countries where it is not publicly provided. Many countries have also experienced increases in house prices which stress traditional systems for financing homeownership. Improving information technology has made simple financial transactions cheaper and easier, but it has also permitted the development of more complex and confusing financial products. The global financial crisis of the late 2000s highlighted the interactions of these trends, and undermined public confidence in the financial system. Finally, within the economics profession the emergence of behavioral economics has opened the eyes of academics to financial behaviors that were not carefully examined before, and to evidence that many households—particularly those with lower income, wealth, and education—are not up to the challenges of managing their financial affairs.

When households lack the intellectual capacity to manage their financial decisions, they make mistakes that lower their own welfare and can also have broader consequences for the economy. For example, if poorer people invest ineffectively and borrow expensively, the wealth of poorer people will grow more slowly than the wealth of richer people even if they have the same savings rates, a problem highlighted by Piketty (2014) and Lusardi, Michaud, and Mitchell (forthcoming). The mistakes of unsophisticated consumers can create rents that distort competition, as emphasized by DellaVigna and Malmendier (2004); Akerlof and Shiller (2015); and Zingales (2015). To attract unsophisticated consumers, financial institutions may lower the up-front costs and raise the hidden costs of financial products, effectively subsidizing consumers who are sophisticated enough to avoid the hidden costs. Sophisticated consumers then have no incentive to adopt easier-to-use financial products, a barrier to constructive financial innovation highlighted by Gabaix and Laibson (2006). Household mistakes that are not purely idiosyncratic but are correlated across households create endogenous risk, which must then be managed by the financial system at considerable resource cost. Important examples include prepayment risk and default risk in mortgages as discussed by Campbell (2006) and Gabaix, Krishnamurthy, and Vigneron (2007), and house
price risk as emphasized by Shiller (2005). Finally, the exploitation of household mistakes by financial institutions can lead to corrosive mistrust of the financial system and the institutions that govern the economy, a problem highlighted by Guiso, Sapienza, and Zingales (2008) and Zingales (2015).

The remainder of this lecture explores the nature of the household finance problem in greater detail, and then discusses policy responses. Although there are numerous traditional arguments for financial regulation of consumer products that do not rely on consumer financial mistakes (Campbell et al. 2011), I concentrate on the more novel and controversial case for intervention in response to such mistakes. In addition, I ignore the important but widely discussed issue of how much households choose to save or borrow, focusing instead on the ways in which they do so.4

Section I summarizes empirical evidence on the broad characteristics of household balance sheets and consumer finance systems in a number of major developed countries. This section shows that there are striking differences across countries which are hard to explain as the result of fundamental differences in tastes or technology. Instead, the structure of household finance appears to be importantly the result of historical social and political choices (some of which may have been made inadvertently). There are also however important common facts across countries. Richer people tend to take more financial risk, and also invest more efficiently. This has implications for the evolution of wealth inequality which are discussed in Section IA.

Section II asks what people find difficult about household finance. This section argues that problems arise not only from traditionally defined financial illiteracy—the failure to understand basic financial concepts—but also from difficulties in understanding the terms of financial contracts, in learning from financial history, and in predicting one’s own behavior and the behavior of others. In other words, people are ignorant not just about accounting, but also about psychology and advanced topics in economics including asset pricing, information economics, and general equilibrium theory. Understood this way, the prevalence of financial ignorance should not be surprising. Consumer financial decisions are complex, and in fact rules of thumb taken from elementary economics often give the wrong answer.

The last two parts of the lecture discuss alternative policy responses to these problems. Educational efforts to improve financial literacy and situation-specific disclosures are both worthwhile, but these alone are unlikely to be sufficient because of the breadth and advanced nature of the knowledge needed to navigate modern financial systems. Hence, there is growing interest in consumer financial regulation. Section III builds on an informal discussion in Campbell et al. (2011) to present an extremely simple welfare analysis of financial regulation in a stylized model with both rational and behavioral consumers.

Section IV considers several specific examples of market trends and policy interventions in consumer finance. This section argues that the Pension Protection Act of 2006 has stimulated favorable trends in the asset allocation of retirement savings, while the CARD Act of 2009 and regulations of the bank overdraft market and the payday lending market have affected the cost structure of short-term unsecured credit. The section also considers a market that appears not to be working well, despite its apparent promise: the market for reverse mortgages. Section V concludes, and an online Appendix provides supporting details.

I. Stylized Facts of Household Finance

Contemporary research on household finance is global in its scope. The comparison of household financial behavior around the world reveals striking differences across countries in average household balance sheets, but also important similarities in the within-country effects of wealth and other household attributes on financial behavior. This section summarizes some key patterns and draws lessons for the field.

Table 1, a refinement of a table in Badarinza, Campbell, and Ramadorai (forthcoming), reports participation rates in various types of assets and liabilities among households in eight leading industrialized countries: the United

---

4 Because of this focus, I do not address the important question of whether households have “present-biased” preferences (Laibson 1997) that lead them to save less than they would if they were able to commit to a savings plan. James Poterba discusses the issue of savings adequacy in his Ely Lecture (Poterba 2014).
Financial assets, nonfinancial assets, and liabilities are listed in descending order of their participation rate among US households. The bottom of the table displays estimates of the participation rates in all risky financial assets (defined to include both bonds and equities) and in equities alone, combining both direct holdings and indirect holdings in mutual funds and retirement accounts. These “look-through” participation rates measure the fraction of households that are exposed to financial risk, regardless of the form in which this exposure is taken.

A number of facts jump out from this table. Deposits and transactions accounts are almost universally held, but there is a minority of “unbanked” households whose size varies enormously across countries. This group is relatively large in the United States. Retirement assets are important in countries with predominantly defined-contribution retirement systems, the Anglo-Saxon and northern European countries. There is a low participation rate in risky assets outside retirement accounts. Directly held bonds play a minor role in household finance in most countries. The look-through participation rates in risky assets and in equities are highly variable across countries, being over twice as large in the Anglo-Saxon countries as in Italy and Spain.

Turning to nonfinancial assets, the homeownership rate is highly variable across countries, particularly low in Germany and high in Spain.

---

**States, Canada, France, Germany, Italy, the Netherlands, Spain, and the United Kingdom.** Financial assets, nonfinancial assets, and liabilities are listed in descending order of their participation rate among US households. The bottom of the table displays estimates of the participation rates in all risky financial assets (defined to include both bonds and equities), and in equities alone, combining both direct holdings and indirect holdings in mutual funds and retirement accounts. These “look-through” participation rates measure the fraction of households that are exposed to financial risk, regardless of the form in which this exposure is taken.

A number of facts jump out from this table. Deposits and transactions accounts are almost universally held, but there is a minority of “unbanked” households whose size varies enormously across countries. This group is relatively large in the United States. Retirement assets are important in countries with predominantly defined-contribution retirement systems, the Anglo-Saxon and northern European countries. There is a low participation rate in risky assets outside retirement accounts. Directly held bonds play a minor role in household finance in most countries. The look-through participation rates in risky assets and in equities are highly variable across countries, being over twice as large in the Anglo-Saxon countries as in Italy and Spain.

Turning to nonfinancial assets, the homeownership rate is highly variable across countries, particularly low in Germany and high in Spain.

---

5 The data sources are the US Survey of Consumer Finances (SCF) (2010); the Canada Survey of Financial Security (2012); the Eurosystem Household Finance and Consumption Survey (HFCS) for Eurozone countries (2008 for Spain, 2009 for the Netherlands, and 2010 for Germany, France, and Italy); and the UK Wealth and Assets Survey (2012). Both the HFCS and the SCF use multiple imputation, but for simplicity we present results for one set of imputations (results are insensitive to this choice). Survey weights are adjusted to impose equal weighting across households. Further details are available in the online Appendix. Christelis, Georgerakos, and Halliasos (2013) conduct a detailed comparison of balance sheets across these countries (with the exception of Canada) and several other European countries, for households over age 50.

6 This look-through calculation, and the analogous calculation in Table 2, require imputation of the asset allocation in retirement accounts using country-specific averages. The online Appendix reports details of this procedure.

---

**Table 1—International Comparison of Participation Rates**

<table>
<thead>
<tr>
<th></th>
<th>USA</th>
<th>Canada</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>Spain</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deposits and transaction accounts</td>
<td>92.6</td>
<td>93.6</td>
<td>99.6</td>
<td>99.0</td>
<td>91.8</td>
<td>94.2</td>
<td>98.1</td>
<td>97.4</td>
</tr>
<tr>
<td>Retirement assets and life insurance</td>
<td>57.6</td>
<td>70.5</td>
<td>37.5</td>
<td>46.5</td>
<td>18.0</td>
<td>49.8</td>
<td>23.6</td>
<td>76.1</td>
</tr>
<tr>
<td>Directly held stocks</td>
<td>15.1</td>
<td>10.0</td>
<td>14.7</td>
<td>10.6</td>
<td>4.6</td>
<td>10.4</td>
<td>10.4</td>
<td>17.5</td>
</tr>
<tr>
<td>Other financial assets</td>
<td>13.4</td>
<td>20.1</td>
<td>10.2</td>
<td>22.0</td>
<td>4.2</td>
<td>10.7</td>
<td>6.8</td>
<td>31.9</td>
</tr>
<tr>
<td>Bonds</td>
<td>13.2</td>
<td>7.4</td>
<td>1.7</td>
<td>5.2</td>
<td>14.6</td>
<td>6.0</td>
<td>1.4</td>
<td>28.2</td>
</tr>
<tr>
<td>Mutual funds</td>
<td>8.7</td>
<td>11.6</td>
<td>10.7</td>
<td>16.9</td>
<td>6.3</td>
<td>17.7</td>
<td>5.6</td>
<td>5.4</td>
</tr>
<tr>
<td>Vehicles and other durables</td>
<td>87.5</td>
<td>100.0</td>
<td>100.0</td>
<td>73.2</td>
<td>95.1</td>
<td>82.6</td>
<td>79.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Main residence</td>
<td>67.2</td>
<td>62.5</td>
<td>55.3</td>
<td>44.2</td>
<td>68.7</td>
<td>57.1</td>
<td>82.7</td>
<td>68.0</td>
</tr>
<tr>
<td>Other real estate</td>
<td>14.4</td>
<td>18.4</td>
<td>24.7</td>
<td>17.8</td>
<td>24.9</td>
<td>6.1</td>
<td>36.2</td>
<td>10.9</td>
</tr>
<tr>
<td>Private businesses</td>
<td>12.1</td>
<td>17.1</td>
<td>10.7</td>
<td>9.4</td>
<td>18.4</td>
<td>4.8</td>
<td>15.0</td>
<td>NA</td>
</tr>
<tr>
<td>Vehicle loans and other debt</td>
<td>48.5</td>
<td>42.7</td>
<td>28.7</td>
<td>21.7</td>
<td>15.3</td>
<td>24.6</td>
<td>27.2</td>
<td>44.0</td>
</tr>
<tr>
<td>Residential mortgage</td>
<td>47.0</td>
<td>33.8</td>
<td>16.9</td>
<td>18.0</td>
<td>9.6</td>
<td>43.9</td>
<td>26.8</td>
<td>36.2</td>
</tr>
<tr>
<td>Credit cards</td>
<td>39.4</td>
<td>24.8</td>
<td>NA</td>
<td>3.4</td>
<td>1.4</td>
<td>4.6</td>
<td>7.3</td>
<td>24.8</td>
</tr>
<tr>
<td>Other mortgage debt</td>
<td>5.4</td>
<td>6.4</td>
<td>10.1</td>
<td>6.0</td>
<td>1.6</td>
<td>2.5</td>
<td>7.3</td>
<td>4.8</td>
</tr>
<tr>
<td>Overdrafts and credit lines</td>
<td>2.1</td>
<td>39.9</td>
<td>7.0</td>
<td>19.8</td>
<td>3.6</td>
<td>20.8</td>
<td>0.6</td>
<td>NA</td>
</tr>
<tr>
<td>Direct and indirect risky assets</td>
<td>61.8</td>
<td>72.7</td>
<td>45.7</td>
<td>54.5</td>
<td>31.4</td>
<td>59.2</td>
<td>30.9</td>
<td>65.2</td>
</tr>
<tr>
<td>Direct and indirect equity</td>
<td>49.1</td>
<td>72.3</td>
<td>45.5</td>
<td>53.5</td>
<td>32.2</td>
<td>58.6</td>
<td>30.3</td>
<td>64.0</td>
</tr>
</tbody>
</table>

Notes: We report the percentages of households which hold respective categories of assets and liabilities. We include accounts with positive balances, as well as ones which households report as active, but which hold amounts equal to zero during the time of the interview. The top part of the table refers to financial and nonfinancial assets and the bottom part to mortgage and non-mortgage debt. Retirement assets include all types of defined contribution plans (public, occupational, or private) which have an account balance. NA denotes asset or debt categories for which holdings are not separately classified, or for which data has not been collected.
On the liability side of the balance sheet, the incidence of mortgage debt is also highly variable across countries but does not line up well with the homeownership rate because southern European housing is much less likely to be mortgage-financed. Auto loans, student loans, and credit card debt are prevalent in Anglo-Saxon countries but not elsewhere. Table 2 reports the average shares of these same assets and liabilities in household balance sheets. Shares are calculated at the household level (including zeros for nonparticipants), and then averaged across households. This procedure weights households equally, whereas summing assets and liabilities across households and then calculating shares gives wealthy households greater weight and can produce quite different numbers. The top part of the table shows the composition of assets, while the lower part shows the composition of liabilities for those households that report liabilities. The very bottom of the table shows the shares of risky assets and equities in household financial assets, looking through retirement accounts to the underlying investments.

The table shows that nonfinancial assets are more important than financial assets in all countries, but particularly so in southern Europe where owner-occupied housing is of overwhelming importance. Among financial assets, retirement assets have the highest share in Anglo-Saxon countries while deposits are more important in European countries. Looking through retirement accounts to the underlying financial investments, the risky shares in financial assets are much larger in Anglo-Saxon countries and to some degree in the Netherlands than elsewhere in Europe. Naturally if one calculated risky asset shares relative to total gross assets, including both financial and nonfinancial assets, the southern European shares would be even lower. On the liabilities side of the balance sheet, Badarinza, Campbell, and Ramadorai (forthcoming) document that households are far more indebted in the Anglo-Saxon countries and the Netherlands than in the other European countries. Putting these facts together, household exposures to financial risks are highly variable across countries.

### Table 2—International Comparison of the Allocation of Household Wealth

<table>
<thead>
<tr>
<th></th>
<th>USA</th>
<th>Canada</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>Spain</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retirement assets and life insurance</td>
<td>13.3</td>
<td>24.1</td>
<td>6.1</td>
<td>10.5</td>
<td>1.5</td>
<td>16.8</td>
<td>1.4</td>
<td>25.1</td>
</tr>
<tr>
<td>Deposits and transaction accounts</td>
<td>11.6</td>
<td>9.9</td>
<td>22.0</td>
<td>30.0</td>
<td>11.9</td>
<td>21.3</td>
<td>10.5</td>
<td>5.9</td>
</tr>
<tr>
<td>Other financial assets</td>
<td>2.1</td>
<td>1.2</td>
<td>1.0</td>
<td>3.5</td>
<td>0.3</td>
<td>0.8</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>Mutual funds</td>
<td>1.3</td>
<td>1.3</td>
<td>0.7</td>
<td>2.4</td>
<td>0.6</td>
<td>1.6</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>Directly held stocks</td>
<td>1.3</td>
<td>1.0</td>
<td>1.0</td>
<td>0.9</td>
<td>0.2</td>
<td>0.6</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>Bonds</td>
<td>0.5</td>
<td>0.2</td>
<td>0.1</td>
<td>0.6</td>
<td>1.6</td>
<td>0.6</td>
<td>0.1</td>
<td>1.1</td>
</tr>
<tr>
<td>Main residence</td>
<td>40.6</td>
<td>31.9</td>
<td>38.9</td>
<td>29.9</td>
<td>53.2</td>
<td>43.3</td>
<td>61.2</td>
<td>34.6</td>
</tr>
<tr>
<td>Vehicles, valuables, and other assets</td>
<td>22.5</td>
<td>22.6</td>
<td>18.5</td>
<td>13.4</td>
<td>19.1</td>
<td>11.1</td>
<td>8.6</td>
<td>28.7</td>
</tr>
<tr>
<td>Private businesses</td>
<td>3.5</td>
<td>2.6</td>
<td>2.7</td>
<td>2.0</td>
<td>3.4</td>
<td>1.7</td>
<td>3.2</td>
<td>NA</td>
</tr>
<tr>
<td>Other real estate</td>
<td>3.2</td>
<td>5.2</td>
<td>9.1</td>
<td>6.7</td>
<td>8.2</td>
<td>2.2</td>
<td>13.3</td>
<td>2.8</td>
</tr>
<tr>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Mortgage debt for primary residence</td>
<td>52.7</td>
<td>38.1</td>
<td>31.5</td>
<td>33.8</td>
<td>35.5</td>
<td>60.6</td>
<td>48.3</td>
<td>51.1</td>
</tr>
<tr>
<td>Vehicle, student loans, and other debt</td>
<td>31.2</td>
<td>28.2</td>
<td>44.5</td>
<td>32.7</td>
<td>50.8</td>
<td>22.5</td>
<td>35.5</td>
<td>34.2</td>
</tr>
<tr>
<td>Credit card debt</td>
<td>12.1</td>
<td>12.4</td>
<td>NA</td>
<td>2.3</td>
<td>3.0</td>
<td>1.1</td>
<td>3.4</td>
<td>9.8</td>
</tr>
<tr>
<td>Other debt secured with real estate</td>
<td>3.3</td>
<td>5.4</td>
<td>15.7</td>
<td>9.5</td>
<td>4.7</td>
<td>2.4</td>
<td>12.2</td>
<td>4.8</td>
</tr>
<tr>
<td>Overdrafts and credit lines</td>
<td>0.7</td>
<td>15.9</td>
<td>8.3</td>
<td>21.7</td>
<td>6.2</td>
<td>13.4</td>
<td>0.5</td>
<td>NA</td>
</tr>
<tr>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Direct and indirect risky assets</td>
<td>38.2</td>
<td>50.3</td>
<td>22.6</td>
<td>29.0</td>
<td>18.5</td>
<td>40.2</td>
<td>15.7</td>
<td>42.2</td>
</tr>
<tr>
<td>Direct and indirect equity</td>
<td>21.2</td>
<td>26.0</td>
<td>9.2</td>
<td>8.6</td>
<td>6.2</td>
<td>16.2</td>
<td>6.2</td>
<td>21.2</td>
</tr>
</tbody>
</table>

Notes: We calculate the respective share of each wealth category relative to the total asset and debt holdings of the household and report averages across the population. The top part of the table refers to financial and nonfinancial assets and the bottom part to mortgage and non-mortgage debt. Retirement assets include all types of defined contribution plans (public, occupational, or private) which have an account balance. NA denotes asset or debt categories for which holdings are not separately classified, or for which data has not been collected.
The balance sheet shares in Table 2 are importantly influenced by the presence of nonparticipants, who have zero shares in certain asset classes. To emphasize this point, Figure 1 plots the average shares of risky assets—including both long-term bonds and equities—across all households in each country (panel A) and across participants in risky asset markets, that is, households with nonzero risky asset shares (panel B). The figure also breaks risky asset holdings into defined-contribution retirement holdings, mutual funds outside retirement accounts, and directly held risky assets. The risky shares in panel A are highly variable across countries, ranging from over 50 percent in Canada to less than 20 percent in Spain. These differences are driven largely by cross-country differences in defined-contribution retirement holdings. Among participants, however, the differences are much smaller, ranging only between 50 percent and 70 percent as shown in panel B. Thus, the main effect of a defined-contribution retirement system is to increase the participation rate in risky financial markets, not the allocation to risky assets conditional upon participation.

The cross-country heterogeneity illustrated in Tables 1 and 2 and Figure 1 implies that household financial systems are strongly influenced by some combination of national financial regulation and national culture in financial matters. While households in each country may take their own system for granted, the world as a whole offers many alternatives—even within this small group of advanced industrial economies—and there is no strong tendency for convergence to a single “ideal” system.7

Within countries, however, the heterogeneity across households displays some important common features. Most notably, there is a strong tendency for the participation rate in risky asset markets to increase with wealth. This is illustrated in panel A of Figure 2, which shows the average risky share by deciles of gross assets in each of the eight countries studied here. An upward slope is visible in all countries, which is steepest in the countries with the highest average risky share. However, there is a much smaller wealth effect on the risky share conditional on participation (panel B). These estimates do not control for covariates of wealth, notably age, but results are broadly robust when one introduces such controls.8

The tendency for richer people to take more financial risk may in part be explained by preferences with declining relative risk aversion, as suggested by Carroll (2002); Calvet, Campbell, and Sodini (2009); Wachter and Yogo (2010);

---

Panel A. All households

Panel B. Risky market participants

**Figure 1. International Comparison of Average Risky Asset (Bond and Equity) Shares**

---

7 Richard Ely would probably not be surprised by this. He wrote of the tendency for regulatory structures to become accepted and hence almost invisible: “We speak continually of the increase of freedom, and imagine often that we have been moving in the direction of no-government. It is probable, however, that laws were never more numerous nor more far-reaching in their consequences than today… They construct a framework within which we willingly move” (Ely 1894, pp. 72–73).

8 Wachter and Yogo (2010) and Guiso and Sodini (2013) present similar analyses and show a greater tendency for risk-taking to increase with wealth among risky market participants. The results here differ for several reasons. First, risky assets are defined to include both stocks and long-term bonds. Second, defined contribution pension assets are included. Third, the cutoffs for wealth deciles are based on the wealth distribution for all households, not just participants. Finally, these results report average, not median, risky shares within each wealth bin. This makes a difference because the cross-sectional distribution of risky shares is right-skewed among households with lower wealth.
and Calvet and Sodini (2014) among others. However, such preferences do not explain why the main channel for the effect is an increase in the participation rate among richer people or why the participation rate remains well below 100 percent even at the top of the wealth distribution. Limited participation among poorer people may be rationalized by fixed costs of participation in risky asset markets (Gomes and Michaelides 2005), but this explanation is less plausible at higher levels of wealth.

A. Investment Strategies and Wealth Inequality

The correlation between wealth and risk-taking is relevant for contemporary concerns about wealth inequality. While the traditional view is that dispersion in wealth at retirement is driven mainly by savings decisions when young (Venti and Wise 2001), attention has shifted more recently to the allocation of savings between riskless and risky assets, and the choice of risky assets. We have already seen that richer people have higher average risky shares in their portfolios. In addition, richer people may invest their risky assets more effectively. Piketty (2014, ch. 12) argues that for both reasons richer people tend to earn higher average returns on their portfolios, and that this is a powerful force increasing wealth inequality.

A difficulty in evaluating this concern empirically is that portfolio returns are difficult to observe. Household survey data are generally inadequate for measuring returns, or even the detailed composition of portfolios within the broad asset classes already discussed. Research on this topic must therefore use the high quality administrative data on all components of wealth that is available in Scandinavian countries, or administrative data on particular types of wealth such as direct stockholdings.

In this part of the lecture I draw on such data to evaluate the potential of investment returns to contribute to wealth inequality. I consider a set of households, indexed by \( i \), with heterogeneous investment strategies. In the absence of income and consumption (and hence in the absence of saving), the evolution of household \( i \)'s wealth is given by \( W_{i,t+1} = W_{i,t}(1 + R_{i,t+1}) \), where \( (1 + R_{i,t+1}) \) is the gross return on household \( i \)'s portfolio. Taking logs,

\[
(1) \quad w_{i,t+1} = w_{i,t} + r_{i,t+1}
\]

\[
= w_{i,t} + E_{t} r_{i,t+1} + \tilde{r}_{i,t+1},
\]

where \( E_{t} r_{i,t+1} \) is the rational (econometrician’s) expectation of the log portfolio return for household \( i \), and \( \tilde{r}_{i,t+1} = r_{i,t+1} - E_{t} r_{i,t+1} \) is the unexpected component of the log portfolio return.

I use the notation \( \text{Var}^{*} \) and \( \text{Cov}^{*} \) to denote the cross-sectional variance and covariance of observations at a point in time. It follows from equation (1) that we can decompose the time-series average change in the cross-sectional variance of log wealth from one period to the next—the average change in wealth inequality—into three terms:

\[9\]He writes “It is perfectly possible that wealthier people obtain higher average returns than less wealthy people… It is easy to see that such a mechanism can automatically lead to a radical divergence in the distribution of capital” (Piketty 2014, p. 430). Lusardi, Michaud, and Mitchell (forthcoming) build a model in which wealth justifies the acquisition of financial knowledge, increasing investment returns and amplifying wealth inequality.

---

**Figure 2. Average Risky Asset (Bond and Equity) Shares across Deciles of Gross Assets**

Panel A. All households

Panel B. Risky market participants

![Graph showing average risky asset shares across deciles of gross assets for different countries.](image-url)
When returns are conditionally log-normal, $E_{t}r_{i,t+1} = \log(E_{t}(1 + r_{i,t+1})) - (1/2)\text{Var}_{t}r_{i,t+1}$. Hence, an uncompensated increase in portfolio variance that leaves the average simple return unchanged lowers the average log return.}

In any one period, the evolution of wealth inequality is also influenced by the cross-sectional covariance between log wealth and expected log returns and the cross-sectional covariance between expected and unexpected log returns. However, these covariances are driven by unexpected common shocks to assets favored by wealthy people, or assets with high expected returns, and thus they average to zero over time. The framework can be extended to include saving as an additional influence on the change in wealth inequality: this adds three new terms which reflect the cross-sectional variance of savings rates and the cross-sectional covariances between wealth and saving, and between saving and expected investment returns. The online Appendix to the lecture presents a more detailed analysis.

There are minor exceptions at the two extremes of the wealth distribution. The very smallest accounts are tiny and tend to hold micro-cap stocks that have delivered high average returns in this period. The largest 0.1 percent of accounts contribute almost 40 percent larger than the sum of the three terms on the right-hand side of equation (2). The same data can be used to calculate the three terms on the right-hand side of equation (2). The sum of the three terms is 0.015 per month, of which a negligible share (less than 0.05 percent) comes from the cross-sectional variance of expected returns, 57 percent comes from the cross-sectional variance of unexpected returns, and 43 percent comes from the cross-sectional covariance between log wealth and expected log returns. The average level of the cross-sectional variance of log account size is 4.55, and the average monthly change in this cross-sectional variance over the sample period is 0.021 per month, almost 40 percent larger than the sum of the three terms on the right-hand-side of equation (2). This implies that inflows to Indian equity accounts, which are not modeled in (2), have increased inequality in equity account size over time. 

To implement this analysis empirically, I report results from two international datasets where both portfolio returns and proxies for wealth can be measured accurately. The first dataset, used in Campbell, Ramadorai, and Ranish (2015), contains a universe of directly held equities from National Securities Depository Limited (NSDL), one of two main electronic securities depositories in India, over the period 2002–2012 at a monthly frequency. All directly held, publicly traded Indian equities are included, and this captures the great majority of risky financial assets for Indian investors, since directly held foreign stocks and mutual funds both have a very low market share in India. Wealth is proxied here by the size of the NSDL equity account. In this dataset, average log returns increase strongly with account size, from 75 to 90 basis points per month among the smallest accounts to almost 140 basis points per month in the top decile of accounts. This pattern reflects generally increasing diversification among larger accounts, which increases average log returns while leaving average simple returns almost flat across the account size distribution. 

The same data can be used to calculate the three terms on the right-hand side of equation (2). The sum of the three terms is 0.015 per month, of which a negligible share (less than 0.05 percent) comes from the cross-sectional variance of expected returns, 57 percent comes from the cross-sectional variance of unexpected returns, and 43 percent comes from the cross-sectional covariance between log wealth and expected log returns. The average level of the cross-sectional variance of log account size is 4.55, and the average monthly change in this cross-sectional variance over the sample period is 0.021 per month, almost 40 percent larger than the sum of the three terms on the right-hand-side of equation (2). This implies that inflows to Indian equity accounts, which are not modeled in (2), have increased inequality in equity account size over time. 

---

(2) $E[\text{Var}^{*}w_{i,t+1} - \text{Var}^{*}(w_{it})]$

$= E[\text{Var}^{*}(E_{t}r_{i,t+1})] + E[\text{Var}^{*}(\tilde{r}_{i,t+1})]$

$+ 2E[\text{Cov}^{*}(w_{it}, E_{t}r_{i,t+1})]$. 

All the returns that appear on the right-hand side of equation (2) are log returns, not simple returns. Cross-sectional variation in expected log returns causes growing cross-sectional dispersion in wealth, consistent with the well-known fact that in an economy without consumption, investors who maximize expected log returns hold the so-called “growth-optimal portfolio” control an ever-increasing share of wealth. 

Hence, differences in diversification across investors can increase wealth inequality through this channel. 

The first term on the right-hand side of (2) is the cross-sectional variance in expected log returns, caused by differences across households in their diversification of uncompensated risk, their willingness to take compensated risk, or their investment skill. The second term is the cross-sectional variance in unexpected log returns, the result of either exposures to idiosyncratic risks or cross-sectionally varying exposures to common shocks. The third term is the cross-sectional covariance between log wealth and expected log returns, which captures any tendency for wealthier households to earn higher expected returns, caused by differences across households, or assets with high expected returns, and thus they average to zero over time. The framework can be extended to include saving as an additional influence on the change in wealth inequality: this adds three new terms which reflect the cross-sectional variance of savings rates and the cross-sectional covariances between wealth and saving, and between saving and expected investment returns. The online Appendix to the lecture presents a more detailed analysis. 

There are minor exceptions at the two extremes of the wealth distribution. The very smallest accounts are tiny and tend to hold micro-cap stocks that have delivered high average returns in this period. The largest 0.1 percent of accounts are somewhat less diversified than smaller accounts, possibly because they belong to corporate executives who are encouraged or required to hold their employer’s stock.
Despite these issues, both datasets imply that the cross-sectional variance of unexpected returns is widely held in Sweden which further lowers the data. In addition, mutual funds are much more operative in the Indian and Swedish data and financial systems between log wealth and expected log portfolio holdings, an important reason for the covariance of unexpected returns in Sweden over the period 2000–2008. Their data are extremely high quality and include all major components of wealth with the exception of private businesses. They measure realized and expected returns on stocks held at the beginning of each year, assuming that no rebalancing occurs during the year, and imposing a simple factor model and the assumption that alphas are zero on average in all wealth groups.

BCS report that the sum of the three terms in equation (2) is 0.062 per year, of which 2 percent comes from the cross-sectional variance of expected returns, 22 percent comes from the cross-sectional variance of unexpected returns, and 76 percent comes from the cross-sectional covariance between log wealth and expected log returns. In the Swedish data the average level of the cross-sectional variance of log wealth is 2.85, and the average annual change in this cross-sectional variance is 0.039, almost 60 percent smaller than the sum of the three terms on the right-hand side of equation (2). Thus, in Sweden unmodeled savings flows have reduced inequality in financial wealth over this period (possibly offset by increasing inequality in housing wealth in a period of rising house prices).

Several important differences between the Indian and Swedish data and financial systems preclude a direct comparison of the results. Since the Swedish data include riskless asset holdings, an important reason for the covariance between log wealth and expected log portfolio returns in Sweden is that, like wealthy people in other countries, wealthier Swedes have higher participation rates and higher average risky asset shares. The presence of riskless assets also lowers the cross-sectional variance of unexpected returns relative to what would be measured in portfolios consisting only of risky assets. These effects are not operative in the Indian data. In addition, mutual funds are much more widely held in Sweden which further lowers the cross-sectional variance of unexpected returns. Despite these issues, both datasets imply that household investment strategies in risky asset markets have substantial effects on wealth inequality relative to directly measured increases in inequality. The cross-sectional variance of expected returns has only a negligible impact, but there are important contributions both from the random realizations of undiversified risky returns, and from the tendency for wealthy people to earn higher average log returns through greater compensated risk-taking and more effective diversification.

Evidence that average log returns vary cross-sectionally is not in itself evidence that households are choosing suboptimal investment strategies. Maximizing average log return is the investment objective only for an investor with log utility: a more risk-averse investor may rationally choose a safer portfolio with a lower average return, while a more risk-tolerant investor may rationally choose a riskier portfolio with a higher simple average return and a lower average log return. However, it cannot be rational to take uncompensated idiosyncratic risk, and such uncompensated risk contributes significantly to the patterns discussed in this section. I now turn to a broader discussion of the evidence that some households make poor financial decisions.

II. What Goes Wrong in Household Finance?

A. Mistakes

Household finance researchers have documented numerous cases where households make decisions that are hard to rationalize using any model of optimal choice. The most extreme cases, although not necessarily the most important ones, have unambiguous optimal decisions that are insensitive to households’ circumstances and preferences. When households fail to make such decisions, their choices are hard to defend. Leading examples in the literature

13 Samuelson (1979), frustrated with the difficulty of convincing investors and academics of this, wrote an article that concludes: “No need to say more. I have made my point. And, but for the last, have done so in words of but one syllable.” The article is a tour de force but is almost incomprehensible because of the difficulty of expressing oneself clearly with such a limited vocabulary. The article also contains a few errors (two-syllable words).

14 One might draw an analogy with tax evasion charges for mob bosses. These are not the most serious crimes, but
include the failure to contribute to 401(k) plans with employer matches and immediate penalty-free withdrawals for older employees (Choi, Laibson, and Madrian 2011), and the failure to locate taxable assets in nontaxable retirement savings accounts (Barber and Odean 2003; Bergstresser and Poterba 2004).15

A related example that is more important for many households, but that is more sensitive to household circumstances, is the failure to refinance a fixed-rate mortgage when it is advantageous to do so. Mortgages are much larger than other financial liabilities for middle-class households, and the savings from refinancing can be substantial. Campbell (2006) observes that many US households fail to refinance even at interest savings that are far above both the thresholds calculated by academics (Agarwal, Driscoll, and Laibson 2013) and the rules of thumb commonly used by financial advisers. In the US, some of this behavior could be explained by unmeasured constraints, but Andersen et al. (2015) show that the same is true in Denmark, where refinancing is always possible even for households with negative home equity or a poor credit score. Similarly, Johnson, Meier, and Toubia (2015) and Keys, Pope, and Pope (forthcoming) document household failures to respond to prequalified refinancing offers.

The existence of such unambiguous mistakes raises the question whether other financial decisions are genuinely the result of rational choice or are also distorted by optimization failures. For example, the failure to participate in the stock market could be justified by large participation costs and high risk aversion, but in many cases is likely to be a mistake. The tendency for apparently suboptimal behaviors to be stronger among households with less income, wealth, and education is circumstantial evidence that supports this interpretation (Campbell 2006; Guiso and Sodini 2013).

B. Five Types of Financial Ignorance

Households that make financial mistakes are almost by definition financially ignorant. But we need greater precision. What exactly is it that consumers fail to understand, and what features of financial problems make them hard for consumers to solve correctly? In the remainder of this section I distinguish five aspects of financial ignorance: ignorance of financial concepts, contract terms, financial history, one’s own behavior, and the strategic behavior of others in market equilibrium. I discuss each of these in turn.

Ignorance of Financial Concepts.—Many households lack understanding of even the most basic concepts needed to solve a financial choice problem. This has come to be called “financial illiteracy.” Hastings, Madrian, and Skimmyhorn (2013) and Lusardi and Mitchell (2014) are two comprehensive recent surveys of the large literature on this subject.

During the past decade, a standard set of questions has been widely used to measure households’ conceptual understanding of finance. Lusardi and Mitchell (2008) formulated three questions, now known as the “Big Three”: one that asks about how interest payments accumulate over time, one on the offsetting effects of inflation and nominal interest rates on real purchasing power, and one on the relative risks of single-stock and mutual-fund investing. The 2009 National Financial Capability Study (NFCS) added two other questions to create the “Big Five”: one on the comparison of 15-year and 30-year mortgages, and one on the relation between interest rates and bond prices. Figure 3 summarizes answers to the Big Five questions given in the 2012 update of the NFCS. The total height of each bar shows the fraction of respondents, broken out by age and gender (the left bars for men and right bars for women), who are financially illiterate as proxied by giving two or fewer correct answers to the five questions.16 Financial illiteracy in this sense is extremely common, affecting an average of 28 percent of men and 44 percent of women across these age groups. It is extremely high among

---

15 Another candidate for this list is the simultaneous holding of both low-interest checking account balances and high-interest credit card debt (Gross and Souleles 2002). However this can potentially be rationalized if households have expenses that can be paid only in cash (Telyukova 2013).

16 This is an extremely low score since the expected number of correct answers is 2.17 if a respondent guesses randomly without answering “Don’t Know” or refusing to answer any questions.
young people: 57 percent of men and 68 percent of women aged 18 to 24 are financially illiterate. This is concerning because many young people do need to make important financial decisions, most obviously with regard to the financing of higher education through student loans. Financial illiteracy diminishes with age during working life, bottoming out at 14 percent for men and 27 percent for women aged 65 to 69, but rises again in old age. This is broadly consistent with a U-shaped life-cycle pattern in various financial mistakes documented by Agarwal et al. (2009), although the bottom of the U is often reached 10 to 15 years earlier in their study. Agarwal et al. attribute the U shape to offsetting effects of accumulating experience and declining cognitive capability with age.

These results are particularly troubling because the Big Five questions are quite easy relative to the level of understanding that is required to make good financial decisions in practice. The first question in particular sets an extremely low bar. It is not about compounding of interest, but about the fact that interest rates are stated on an annual basis so that interest is received five times over five years.

A person could answer this question correctly without any understanding of compounding. Stango and Zinman (2009) report evidence that consumers fail to understand compounding and seem to think about interest rates as if they were simple rates, charged or paid on a constant base, rather than compound rates, charged or paid on a base that changes over time. This "exponential growth bias" leads consumers to underestimate the future value of an amount saved, given an interest rate and time horizon, and to underestimate the interest rate implied by a given stream of level payments, maturity, and principal amount of a loan. Exponential growth bias makes it hard for consumers to compare up-front costs with interest costs, as in the case of a mortgage whose interest rate can be reduced via the payment of initial points.

Survey respondents have unusual difficulty answering the question on risk. While the proportion of incorrect answers is no higher than for some other questions, the proportion of correct answers is lower and a large fraction of respondents say they don’t know the answer. This matters because an understanding of diversification and the trade-off between risk and return is necessary to make an informed decision about asset allocation in retirement saving, and essential for effective investing in risky asset markets as discussed in the previous section of this lecture.

The Big Five questions on mortgages and bond prices are also elementary compared to the advanced conceptual knowledge that is required to select and manage a mortgage. Campbell and Cocco (2003) discuss the subtle risk considerations that are involved in choosing a fixed-rate versus an adjustable-rate mortgage. Once a household has a fixed-rate mortgage, the optimal refinancing strategy is the solution to a difficult real options problem whose parameters include the fixed costs of refinancing, the expected remaining life of the mortgage (which depends on a household’s moving probability), the tax treatment of mortgage interest, and the volatility of interest rates. Agarwal, Driscoll, and Laibson (2013) have only recently published a closed-form solution for this problem. Rules of thumb used by financial planners and

---

17 This question reads: Suppose you had $100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow: [more than $102; exactly $102; less than $102; do not know; refuse to answer]

18 This question reads: Do you think that the following statement is true or false? “Buying a single company stock usually provides a safer return than a stock mutual fund.” [true; false; do not know; refuse to answer]
popularized in financial advice handbooks provide only rough approximations to the optimal solution. The decision to default on an underwater mortgage is equally challenging (Campbell and Cocco 2015).

Despite these limitations of the Big Five financial literacy questions, they remain popular because they are brief and straightforward—and hence easy to insert into general surveys—and because responses can be compared with many other surveys that have been conducted around the world.\(^{19}\) Higher scores on these questions, and particularly the question about risk, correlate with positive answers to questions about saving and retirement planning (Lusardi and Mitchell 2011, 2014). While this does not establish causality, it is plausible that a poor understanding of investment alternatives makes it harder to formulate a financial plan.

Low financial literacy scores are associated with low income, wealth, and education, implying that financial illiteracy—like suboptimal investment behavior—is a particular problem among people with low socioeconomic status. While it may be optimal for poorer people to invest less in financial literacy, as pointed out by Lusardi, Michaud, and Mitchell (forthcoming), this raises concerns that a complex financial system particularly disadvantages poorer people and contributes to wealth inequality.

**Ignorance of Contract Terms.**—The difficulties people have with household finance are not just the result of innumeracy. There is also evidence that people don’t understand the terms of financial contracts, as would be the case if they fail to read the “small print.” For example, Bucks and Pence (2008) show that collectively, mortgage borrowers with adjustable-rate mortgages (ARM) underestimate the extent to which their mortgage rate can change. Lee and Hogarth (1999) report evidence that mortgage borrowers fail to understand the fees that create a wedge between the contract interest rate on a mortgage and the annual percentage rate (APR). Stango and Zinman (2014) show that consumers are often unaware of the circumstances that trigger bank overdraft fees. Similar concerns with regard to credit card fees were an important motivation for the CARD Act of 2009.

Some consumers appear to be unaware of the options that are built into certain financial products. The refinancing option in a fixed-rate mortgage is a particularly important example, as discussed above. In an extreme case, even letters reminding borrowers of a preapproved refinancing opportunity do not trigger action, possibly because borrowers are suspicious that the offer is “too good to be true” (Johnson, Meier, and Toubia 2015; and Keys, Pope, and Pope forthcoming).

When households fail to understand all the costs of a financial product, firms have an incentive to lower salient “front-end” costs and increase obscure “back-end” costs. Thus, the complexity of financial products may be intentional as suggested by DellaVigna and Malmendier (2004); Célérier and Vallée (2015); and Grubb (2015) among others. If sophisticated consumers can avoid back-end costs by altering their behavior (as in the case of credit card or overdraft fees), competition may result in lower prices to sophisticated consumers who thus receive a cross-subsidy from unsophisticated consumers (Gabaix and Laïbson 2006). I discuss this issue further in Section III.

**Ignorance of Financial History.**—In order to make savings and asset allocation decisions, people need to form views about the likely returns on alternative investment strategies. It is rational to do this using all the historical data that are available. However, it appears that in many cases, people rely on their own much more limited and specific experiences to form their beliefs.

One type of evidence shows that people react to experiences shared within their cohort, for example of stock returns and inflation (Malmendier and Nagel 2011, 2016).\(^{20}\) Another type of evidence looks at idiosyncratic experiences that vary cross-sectionally within a cohort. For example, Choi et al. (2009) show that households save more in response to high idiosyncratic

---

\(^{19}\) Some attempts are being made to assess conceptual financial knowledge more thoroughly, for example, in van Rooij, Lusardi, and Alessie (2011) and in the OECD’s Programme for International Student Assessment (PISA) which works with high-school students (OECD 2014).

\(^{20}\) A related strand of the behavioral finance literature (for example, Barberis et al. 2015) argues that investors tend to extrapolate aggregate returns that have been realized over the recent past. Here too the problem is that recent data are overweighted relative to longer-term financial history.
returns in their 401(k) accounts, consistent with the idea that they extrapolate these returns into the future. Campbell, Ramadorai, and Ranish (2015) and Huang (2015) show that households increase their investment activity and portfolio tilts to equity styles and industries, respectively, in which they have experienced high returns on the particular stocks they picked within those styles and industries.

In principle it is possible that responses to idiosyncratic investment experiences could result from investors learning about their skill, either at 401(k) investing in general, or at stock-picking within specific styles and industries. However, several recent papers have documented responses that cannot plausibly be attributed to learning about skill. Knüpfer, Rantapuska, and Sarvimäki (2014) show that heterogeneous labor market experiences during the Finnish Great Depression of the early 1990s have long-lasting impacts on Finnish households’ willingness to take financial risk, and Anagol, Balasubramaniam, and Ramadorai (2015) document that the random allocation of IPO shares to Indian households affects their subsequent equity investment behavior.

A related phenomenon is the tendency of individual investors to overweight familiar assets, in effect treating them as safer than unfamiliar ones. This tendency generates “home bias” or overweighting of domestic stocks relative to international ones (Lewis 1999; Bekaert et al. 2015); overweighting of local and culturally familiar stocks (Grinblatt and Keloharju 2001; Huberman 2001); overweighting of stocks in an investor’s own industry (Massa and Simonov 2006); and even the extreme case of overweighting the stock of the investor’s employer (Benartzi 2001). Some papers have attempted to justify such bias as the result of informational advantage (Massa and Simonov 2006; Van Nieuwerburgh and Veldkamp 2009), but evidence on the performance of overweighted familiar stocks is mixed (Benartzi 2001; Døskeland and Hvide 2011), while there is some evidence that home bias diminishes with education and financial literacy (Bekaert et al. 2015).

Ignorance of Self.—An important question is whether people understand their own difficulties with financial problem-solving. The same surveys that administer financial literacy questions often ask respondents to assess their own financial knowledge. Figure 3 breaks the financially illiterate population into two groups: those who express high confidence in their financial knowledge (giving themselves at least five out of seven possible points), and those who do not. The dark shaded portion of each bar represents self-confident financial illiteracy. Remarkably, this is more common than financial illiteracy with moderate or low confidence, in all age groups and both genders. The self-confident share of financially illiterate individuals is higher among men, and it increases with age. More generally, seniors retain their average confidence even while their average financial literacy scores decline with age.

At the individual level, one can detect a positive correlation between financial literacy and self-confidence, but the relationship is quite weak. Regressions of the confidence score on dummies for correct answers to each of the Big Five questions in the 2012 NFCS deliver statistically significant and positive coefficients, but the $R^2$ of the relationship is below 6 percent both in the population as a whole, and among those under 65 and over 65 years old. While measurement error can certainly reduce the measured correlation between financial literacy and confidence, it does appear that people are heterogeneous in both dimensions.

These findings imply that there are two groups of people we should be concerned about. Financially illiterate people with low self-confidence are likely to have a strong demand for assistance with financial problems, and to be relatively responsive to defaults and other nudges. Overconfident financial illiterates with high self-confidence may make different types of mistakes, such as underdiversifying or trading excessively (Barber and Odean 2000), and these mistakes may be harder to correct.

A related issue is that many financial problems require people to predict their own future behavior. In some cases, this may be relatively straightforward, as when families estimate the length of time they will occupy a home while choosing a mortgage. In other cases, it may be challenging, as with financial instruments that require attention to manage correctly. Some

---

21 Bhamra and Uppal (2014) present a formal model of this behavior and show that it can have sizable negative effects on household welfare.
people may fail to predict their own inattention to future bank account balances (in the case of overdraft fees) or interest rates (in the case of mortgage refinancing). Grubb (2015) surveys a growing literature on this type of overconfidence.22

Ignorance of Incentives, Strategy, and Equilibrium.—A final aspect of financial ignorance is naïveté, a failure to understand the incentives faced by other parties and the effects these incentives have on their strategic behavior and on market equilibrium. There is direct evidence that this is a common phenomenon. A report to the European Commission by Chater, Huck, and Inderst (2010, Table 4.50), for example, surveys European purchasers of retail financial products and finds that over 40 percent of them are unaware of any payments made to financial product providers, salespeople, brokers, or financial advisers.

The behavior of financial market participants also provides strong if circumstantial evidence that households are naïve about incentive effects. Malmendier and Shanthikumar (2007), for example, study the stock recommendations of sell-side security analysts, which are well known to be biased upward, and show that small equity traders respond as if they take the recommendations literally while large traders appear to adjust for the bias. Audit studies of US financial advisers (Mullainathan, Noeth, and Schoar 2012) and Indian life insurance agents (Anagol, Cole, and Sarkar 2012) find that these advisers tend to recommend products that improve their own compensation over the most suitable and affordable products for clients. Such behavior has greater potential to increase advisers’ income if their clients are unaware of it.

A series of papers studying mutual fund performance has shown that actively managed funds sold through brokers, particularly domestic equity funds, tend to underperform both directly sold mutual funds and broker-sold index funds (Bergstresser, Chalmers, and Tufano 2009; Del Guercio and Reuter 2014; Reuter 2015). This is true even before broker sales fees, and all the more so after fees. Yet index funds have a market share below 3 percent among broker-sold funds, and mutual funds’ payments to brokers predict their inflows positively and their performance negatively (Christofferson, Evans, and Musto 2013).

Finally, the degree of price dispersion in financial products such as index mutual funds (Hortaçsu and Syverson 2004), mortgages (Woodward and Hall 2012), and credit cards (Stango and Zinman 2016) implies a very high return to time spent shopping and negotiating in retail financial markets. Yet it appears that many households do minimal shopping (e.g., Chater, Huck, and Inderst 2010, Table 4.65), a fact which helps to explain the survival of high-cost providers.

A closely related type of financial ignorance is the failure to understand how other market participants respond to relevant information that they possess. The theory of rational expectations equilibrium assumes that people correctly infer the information of others from market prices, but such inference is extremely challenging. Behavioral finance models from Miller (1977) to Eyster, Rabin, and Vayanos (2015) and Glaeser and Nathanson (2015) show that market price formation can be radically different when this inference is absent or imperfect. Even when prices are unaffected (perhaps because there are sophisticated, well capitalized arbitrageurs in the market), naïve investors will suffer poor performance if they are unable to learn rationally from market prices.

C. Easy Fixes? Financial Education and Disclosures

A natural response to financial ignorance is to try to cure it through financial education or disclosures. These are variants of the same strategy, empowering individuals to understand financial problems better and thereby to make better decisions. The difference is that financial education is more general, particularly when offered early in life, while disclosures are specific to particular financial problems.

The effectiveness of financial education can only be measured if one can find a source of exogenous variation in educational exposure. Bernheim, Garrett, and Maki (2001) pioneered the use of state-level variation in high school financial curriculum mandates, and reported some evidence in survey data that financial

---

22 An important theme of this literature, outside the scope of this lecture, is that consumers overestimate the self-control they will have over their own future consumption spending (Heidhues and Köszegi 2010).

The provision of unbiased financial advice can be understood as a more specific and outcomes-oriented version of financial education. In order to be effective, financial advice must be of high quality and households must be willing to follow it. Foerster et al. (forthcoming) studies a short personal finance course offered by the US Army to new enlistees and documents some positive results for debt management. Hastings, Madrian, and Skimmyhorn (2013) and Lusardi and Mitchell (2014) survey this literature in greater detail, while Willis (2008) offers a critical perspective.

The reluctance of households to follow advice may limit the quality of the advice they receive. The audit studies of Mullainathan, Noeth, and Schoar (2012) and Anagol, Cole, and Sarkar (2012) find that advisers rarely argue against the expressed beliefs of prospective clients, regardless of the merits of these beliefs. This is consistent with the view that advisers fear losing business if they contradict clients. Public provision of unbiased advice, as offered by the UK government’s Money Advice Service (https://www.moneyadvice.service.org.uk/en) is one possible response to this problem. However it is hard to customize such advice to households’ specific circumstances. Inderst and Ottaviani (2012) survey the literature on financial advice, emphasizing the effects of consumer naiveté on market equilibrium.

Putting these results together, while economics educators find it natural to call for increased financial education, the evidence for its effectiveness is quite limited. And one must also keep in mind that financial education is costly because it competes with other educational priorities in high school, and with productive hours in the workplace. In my view financial education is desirable but cannot be the whole solution to the problem of financial ignorance.

Disclosures are a popular alternative because they target specific decisions that people find hard to make. Disclosure requirements are familiar from other contexts, such as food labeling, and have modest costs compared to other types of financial regulation. There is evidence that disclosures can alter consumer financial behavior, particularly when they are designed with some psychological sophistication. Bertrand and Morse (2011), for example, show that disclosures highlighting the dollar cost of repeated payday borrowing reduce the number and size of payday loans subsequently taken out.

Disclosures however have several limitations. First and most obvious, disclosures cannot work if consumers are too financially ignorant to understand them. Second, it is sometimes possible for financial service providers to neutralize the effects of disclosures by highlighting other information that distracts consumers. Ru and Schoar (2015) highlight this problem in the context of credit card marketing. Third, if disclosure requirements are not very carefully designed, they can be circumvented by offering alternative products or by increasing alternative costs that are not covered by the requirements. Anagol and Kim (2012) and Duarte and Hastings (2012) present examples of such cost-shifting in the Indian and Mexican asset management industries.

Given these limitations of financial education and disclosures, it is natural to consider more stringent interventions in consumer financial markets that alter consumer behavior more directly. The theoretical analysis of such interventions is the subject of the next section.

III. The Behavioral Welfare Economics of Consumer Financial Regulation

A. A Simple Model of Costly Regulation

Behavioral welfare economics analyzes the welfare implications of intervention in markets where some consumers are making mistakes and failing to maximize their true utility. The social planner is assumed to know agents’ true utility and to be at least partly paternalistic, that is, the planner places some weight on agents’ true utility as opposed to their self-perceived

23 The reluctance of households to follow advice may limit the quality of the advice they receive. The audit studies of Mullainathan, Noeth, and Schoar (2012) and Anagol, Cole, and Sarkar (2012) find that advisers rarely argue against the expressed beliefs of prospective clients, regardless of the merits of these beliefs. This is consistent with the view that advisers fear losing business if they contradict clients. Public provision of unbiased advice, as offered by the UK government’s Money Advice Service (https://www.moneyadvice.service.org.uk/en) is one possible response to this problem. However it is hard to customize such advice to households’ specific circumstances. Inderst and Ottaviani (2012) survey the literature on financial advice, emphasizing the effects of consumer naiveté on market equilibrium.

24 Bordalo, Gennaioli, and Shleifer (2015) point out that misleading marketing can sometimes take the form of reminding consumers of unlikely but salient events, such as breakage of a new television set that induces purchase of overpriced insurance. In this case a disclosure policy would have the difficult task of focusing consumers on information about breakage rates and actuarial costs.
utility. In this situation the planner may be able to increase social welfare by intervening to discourage mistaken choices.

Thaler and Sunstein (2008) use the word “nudge” to describe an intervention that affects the decisions of behavioral consumers while imposing negligible costs on rational consumers. The canonical example is a default contribution rate or asset allocation in a retirement savings plan, which is accepted by many participants even though any participant can opt out by making an alternative choice. Nudges are appealing because of their minimal impact on rational consumers, but they have limitations. Overconfident behavioral agents may opt out at high rates, particularly when offered defaults that are far from their unconstrained choices—that is, particularly in cases where they are making large mistakes. It can be hard to design defaults that are sufficiently customized to accommodate heterogeneity in optimal choices, and that adapt to changing household circumstances over time. Financial service providers have incentives to circumvent defaults, just as they do disclosures, by making it easy or tempting to opt out.

If nudges are unavailable or inadequate, consumer financial regulation must confront the trade-off between the benefits of intervention to behavioral agents, and the costs to rational agents. In this section I present a model to illustrate this trade-off. The model is extremely simple and neglects any interactions between behavioral reasons for consumer financial regulation and more traditional market failures (Haltiwanger and Waldman 1989; Campbell et al. 2011; Jackson and Rothstein 2015). The model also ignores the concern, important in practice, that financial product providers can circumvent regulation or influence it through political pressure. Much more general treatments of behavioral welfare economics, extending the canonical public finance literature, can be found in O’Donoghue and Rabin (2006); Mullainathan, Schwartzstein, and Congdon (2012); and Farhi and Gabaix (2015).

Two financial products are available to consumers, A and B. Product A, the “plain-vanilla” product, delivers utility normalized to zero for all consumers: \( U_A = 0 \). Product B, the “problematic” product, is correctly preferred by some consumers but mistakenly chosen by others. As motivating examples, product A could be a bank account without overdraft protection while product B could be the same account with overdraft protection; or product A could be informal lending within the family while product B could be a payday loan.

A continuum of agents, with mass normalized to 1, believe that they receive nonnegative utility \( U_B = u \) from product \( B \) and \( \frac{u}{h} \) from product \( A \) varies across agents with probability density function \( f(u) \), cumulative distribution function \( F(u) \), and upper bound \( h \), but each agent knows her own value of \( u \). At each level of \( u \), a fraction \( b \) of the agents are behavioral agents who receive lower utility from the product, \( U_B = u - 1 \). In other words these agents overstate the utility they receive from product B by a fixed amount normalized to unity. The overstatement might be due to a mistake in assessing the suitability of the product ex ante, or a mistake in managing it ex post. The remaining fraction \((1 - b)\) of the agents are rational and receive utility \( u \) as they expect. Behavioral and rational agents are identical in all other respects, and this rules out the possibility of nudging behavioral agents without impacting rational agents, or of imposing “qualified investor” rules to exclude only behavioral agents from the product-B market.

For a simple parametric example, assume a uniform distribution for \( u \) between 0 (the lower bound consistent with choosing product B) and \( h \) (the upper bound). The parameter \( h \) captures the heterogeneity of agents, both within the B-choosing group and between the average of that group and the unmodeled A-choosing group. With this specification, \( f(u) = 1/h \) and

---

25 The assumption that true utility is known is not innocuous (Goldin 2015). It is particularly problematic in cases where agents maximize expected utility of random outcomes and disagree about state probabilities. Brunnermeier, Simsek, and Xiong (2014) propose a welfare criterion that can be used in this case.

26 Examples of traditional market failures that may be worsened by behavioral biases are market power in search models, which may be increased by households’ reluctance to search, and negative externalities such as the negative effects of debt collection on people who do not owe money (a major source of complaints to the CFPB). It is also possible to write down models in which behavioral biases ameliorate traditional market failures and, in the extreme, even increase social welfare.

27 There can also be agents who prefer product A in the absence of regulation, but these agents are unaffected by regulation and can be dropped from further consideration.
we have $F(c) = c/h$, and a regulatory charge equal to $h$ is sufficient to close down the whole market for product $B$. The social welfare function in this example can be written for $c \leq h$ as

$$W(c) = \left(1 - \frac{c}{h}\right)\left[\left(\frac{h + c}{2}\right) - b\right].$$

The derivative of social welfare with respect to the regulatory charge is

$$\frac{dW}{dc} = \frac{b - c}{h},$$

which is positive at $c = 0$ for the usual reason that rational agents affected by the regulation are indifferent at the initial optimum, while behavioral agents gain (O’Donoghue and Rabin 2006; Dávila 2015). The second derivative is $-1/h$, so the second-order condition holds. The negative second derivative results from the fact that as $c$ increases, the rational agents who respond at the margin are increasingly disadvantaged by the switch from product $B$ to product $A$, while the gain to the marginal behavioral agents is declining. Even though each behavioral agent overstates the utility of product $B$ by the same amount, the self-perceived utility of product $B$ relative to product $A$ is greater, and the true disutility is smaller, for those behavioral agents who respond at the margin to higher levels of $c$.

The optimum policy sets $dW/dc = 0$ provided that $c \leq h$, which implies that the optimum $c^*$ is

$$c^* = b$$

for $b \leq h$. Such a charge induces a fraction $b/h$ of the initially $B$-choosing agents to switch to product $A$. If $b > h$, then any charge above $h$ is sufficient to close down the market for product $B$ and achieve the social optimum.

These results illustrate a point emphasized by Campbell et al. (2011), that product-level financial regulation should be stronger when there are more behavioral agents in the market and when the value of a product to rational agents is smaller relative to the mistakes made by behavioral agents.\[28\] As $h$ becomes small, product $B$...
offers little even to rational agents, and the prevention of mistakes dominates all other considerations. This is the premise of Senator Elizabeth Warren’s famous analogy between mortgages and toasters: consumers all require the same features in toasters, and we take it for granted that toaster design is strictly regulated for safety. With a large $h$, however, there is a meaningful trade-off between the protection of consumers who mistakenly choose product $B$ and the benefits that rational consumers obtain from the product.

The analysis above can be adapted to allow the regulator to place some weight on agents’ self-perceived utility. This might reflect uncertainty about the true utility function, or antipaternalistic political pressure on the regulator. Suppose that the regulator places a positive weight $\beta < 1$ on the agents’ true utility, and weight $1 - \beta$ on the agents’ self-perceived utility, that is, the regulator is only paternalistic to the degree $\beta$. Define a parameter $b^* = \beta b$. Then the previous expressions for the social welfare function and the optimal policy hold after substituting $b^*$ for $b$. As the regulator becomes less paternalistic, $\beta$ and $b^*$ decline, and it is as if there are more rational agents and fewer behavioral agents.

Distributional considerations push in the opposite direction. If behavioral agents have lower wealth and income, the social planner may place greater weight on gains to them than on losses to rational agents. Within the simple linear modeling framework used here, this is equivalent to the social planner setting $b^* > b$, acting as if there are more behavioral agents and fewer rational agents.

*What if the Regulatory Charge Cannot Be Fully Rebated?*—In many consumer regulatory applications, it is not realistic to assume that interventions take the form of taxes delivering revenue that can be rebated to consumers. Instead, the deadweight burden of financial regulation is often emphasized by its opponents. Suppose that a fraction $\alpha$ of the regulatory charge $c$ cannot be rebated but instead is a deadweight cost. Then

$$W(c) = \int_c^h [u - b - \alpha c] f(u) du.$$  

In the parametric example,

$$W(c) = \left(1 - \frac{c}{h}\right) \left[\left(\frac{h + c}{2}\right) - b - \alpha c\right],$$

and

$$\frac{dW}{dc} = \left(\frac{1}{h}\right)\left[b - \alpha h - (1 - 2\alpha)c\right].$$

For the problem to be well behaved, we need a small $\alpha$ such that $\alpha < 1/2$ and $\alpha < b/h$. If these conditions are met, the optimal policy is

$$c^* = \frac{b - \alpha h}{1 - 2\alpha},$$

provided that $c^* < h$. If $c^* \geq h$, any charge above $h$ is sufficient to close down the market for product $B$ and achieve the social optimum.

It is easy to see from equation (10) that the scale of the optimal intervention is increasing in the fraction of behavioral agents $b$, and decreasing in $h$ whenever the deadweight cost of intervention is positive. However the optimal intervention in the presence of a deadweight cost need not be smaller than the optimal intervention in the previous case without a deadweight cost. For example, if $b = h/2$ then $c^* = h/2$ regardless of the value of $\alpha$. And in general the derivative $dc^*/d\alpha$ has the same sign as $(c^* - h/2)$. The reason is that increasing the regulatory charge $c$ shrinks the number of people who pay the charge, and beyond a certain point further increases in $c$ reduce the total deadweight loss. Given the parametric assumptions we have made the greatest deadweight loss is at $c^* = h/2$, corresponding to an intervention that cuts the size of the product-$B$ market in half. To the right of this point, the optimal response to an increase in the deadweight cost of intervention is to increase the scale of the intervention to reduce the number of people paying the charge.
Of course, the social gain from any given intervention does decline in $\alpha$. If $\alpha$ is large enough, then the problem ceases to be convex. For example, if $\alpha = 1$ then

\begin{equation}
\frac{dW}{dc} = \left( \frac{b - h}{h} \right) + c.
\end{equation}

In this case, when $b < h$ we have $dW/dc < 0$ at the point $c = 0$; and the second derivative is always positive. The optimal policy is to set $c = 0$ or $c = h$. Equation (8) implies that $W(0) = h/2 - b$, while $W(h) = 0$. Thus, the optimal policy is $c = h$ if $b \geq h/2$, and $c = 0$ otherwise. In this case the regulator will shut down the market altogether if in the costless case it would reduce market size by more than half, and will avoid any intervention otherwise.

C. Intervention against Mistakes

The interventions considered so far are crude in the sense that they penalize a product that is sometimes used mistakenly, by imposing a regulatory charge, but do not specifically target the mistakes made by behavioral agents. I now consider interventions that do target such mistakes.

Suppose that product $B$ is not inherently inappropriate for behavioral agents, but requires some sophistication to use properly. The consumer financial regulator can make the product easier to manage, thereby reducing the scale of behavioral mistakes, but the cost of such a requirement is borne by all users of the product. In this case the size of the behavioral mistake may be written $g(c)$, where $g(0) = 1$, and $g(c) \geq 0$ and $g'(c) \leq 0$ for all $c$. Social welfare becomes

\begin{equation}
W(c) = \int_c^h \left[ u - bg(c) - \alpha c \right] f(u)du,
\end{equation}

where as before $\alpha$ is the deadweight cost of the intervention. It is natural to assume that $\alpha = 1$ in this case, since real resources are being used to improve the ease of use of product $B$.

For a simple parametric example, consider the previous case of a uniform distribution for $u$ and assume $g(c) = 1 - \theta c$ for $c \leq 1/\theta$. Then the social welfare function is

\begin{equation}
W(c) = \left( 1 - \frac{c}{h} \right) \left[ \left( \frac{h + c}{2} \right) - b - \alpha^* c \right],
\end{equation}

where $\alpha^* = \alpha - b\theta$. This has the same form as equation (9), but with a parameter $\alpha^* < \alpha$. Previous results therefore apply substituting $\alpha^*$ for $\alpha$.

An important difference from the previous analysis is that now it is possible for $\alpha^*$ to be negative. If this is the case, then consumer heterogeneity (which also measures the average benefit of product $B$ to rational agents) increases the scale of the optimal intervention $c^*$, that is $dc^*/dh > 0$. The reason is that a high $h$ makes the financial product valuable to consumers who know how to manage it properly, so it is worthwhile to devote resources to mitigating the mistakes of behavioral agents in using the product.

The effectiveness of the intervention in mitigating mistakes is measured by the parameter $\theta$. It is not necessarily the case that greater effectiveness increases the optimal scale of the intervention, for reasons similar to those discussed earlier. However, greater effectiveness does increase the social benefit of any given intervention.

Mistakes and Transfers.—Some mistakes by behavioral agents are product management failures that generate fees to product providers. Examples might include credit card late payment fees, or bank overdraft fees. Gabaix and Laibson (2006) emphasize that in a competitive financial market, fees from behavioral agents are passed on to all purchasers of the product in the form of lower product prices. Thus, in equilibrium rational agents benefit at the expense of behavioral agents. Such a mechanism potentially affects the welfare analysis of consumer financial regulation because losses to behavioral agents are transfers to rational agents, not social losses.

To understand the issue, I now alter the simple product choice model so that each behavioral agent generates a unit fee that is passed on to all purchasers of product $B$. This is the extreme case where the misperceptions of behavioral agents are entirely about transfers that they make to rational agents. The subsidy per purchaser is $b$, and if the subsidy is removed the price of product $B$ increases by $b$ so rational agents’ net-of-cost utility declines by this amount.

Interventions against products, of the sort discussed in Section IIIIB, have the same social benefits in this case as they did before, and accordingly the optimal scale of these interventions is unchanged. The reason is that
Interventions against products affect rational and behavioral agents in the same proportion that they exist in the population. In effect the social planner uses the preferences of a weighted average of rational and behavioral agents, and transfers net out from this weighted average.

Interventions against mistakes, of the sort discussed in this section, have lower benefits and a smaller optimal scale in the presence of transfers. Eliminating mistakes benefits behavioral agents but, if there are transfers, it hurts rational agents and the social planner must take this into account.

To illustrate this point, consider an intervention with cost $c$ that reduces the size of the behavioral agents’ mistakes to $g(c) = 1 - \theta c$. Such an intervention reduces the subsidy to purchasers of the product by $b\theta c$, raising its price accordingly. Only agents whose self-perceived utility $u > c(1 + b\theta) = \hat{c}$ will now buy the product, so the intervention has an amplified effect on the size of the market for product $B$. Define a parameter $\hat{\alpha} = \alpha/(1 + b\theta)$. Then the social welfare function takes the same form as equations (7) and (8), substituting $\hat{c}$ for $c$ and $\hat{\alpha}$ for $\alpha$. There is no direct effect of $g(c)$ on social welfare, as there was in equation (12), because transfers cancel out of the social welfare function. The optimal intervention is generally smaller than in the case without transfers, because an intervention of any given scale has an amplified effect on the size of the market for product $B$.

IV. Recent Developments in Consumer Financial Markets

This lecture has made the case that consumer financial illiteracy is a sufficiently serious problem, and hard enough to cure through financial education and disclosures, that some degree of financial regulation can be justified even if it imposes costs on rational households. The previous section developed this point in a simple theoretical framework with a regulator using a paternalistic social welfare function. In this section I review some developments in consumer financial markets over the past decade and interpret them in this light. In turn I consider trends in the asset allocation of retirement saving, in fees charged for unsecured consumer credit through credit cards, bank overdrafts, and payday loans, and in the market for reverse mortgages.

A. Asset Allocation in Retirement Saving

The US retirement system is a particularly clear illustration of the general truth that household financial decisions occur in an institutional context and are importantly shaped by that context. The retirement system in place today is the result of two trends, one operating over the last 40 years and the other over the last 15 years. The first trend is the gradual shift from defined-benefit (DB) to defined-contribution (DC) retirement plans, a shift inadvertently initiated by the Employee Retirement Income Security Act (ERISA) of 1974 and the creation of tax-favored 401(k) plans a few years later, but also driven by long-term economic forces such as increasing life expectancy which make DB pension promises more expensive for private employers. This trend has forced individuals to take greater responsibility for their own retirement planning. The second trend, in part a response to the first, is the increasing use of “choice architecture” (Thaler and Sunstein 2008) to guide individuals’ choices through defaults and similar mechanisms. Behavioral economists have documented the power of 401(k) plan defaults to affect enrollment, contribution rates, and asset allocation (Madrian and Shea 2001; Choi et al. 2004; Beshears et al. 2009; Mitchell and Utkus 2012). These results are striking because plan participants can opt out of defaults at low cost, thus defaults are low-intensity interventions with only modest side effects on rational households.

In the specific context of asset allocation, the US regulatory system changed in the mid-2000s with the Pension Protection Act (PPA) of 2006 and subsequent rule-making by the Department of Labor. Prior to the PPA, employers hesitated to use risky funds as default investments for fear of bearing legal responsibility for any losses incurred. This encouraged the use of low-return money market funds for retirement saving. Under the PPA, certain categories of broadly diversified funds are qualified default investment alternatives (QDIA) that protect employers from any responsibility for loss. QDIAs include balanced funds (with risk levels suitable for the company’s workforce as a whole), professionally managed accounts tailored to an individual’s preferences and situation, and target-date funds (TDF). The latter are a cheaper alternative to professional asset management with
risk levels that are calibrated to an individual’s retirement date and that gradually decline over working life.\textsuperscript{31}

The growth of TDFs is a striking phenomenon of the last ten years. Mitchell and Utkus (2012), using data on a panel of 401(k) plans administered by Vanguard, show that TDFs have been adopted both through their designation as defaults in 401(k) plans, and through active choice by plan participants. Even a purely administrative default designation, without any auto-enrollment, increases TDF adoption, consistent with the view that participants have an implicit demand for investment advice and pick plans that they regard as recommended by their employers. TDF adoption correlates with broader changes in 401(k) portfolios (notably an increase in the share of diversified equity among younger investors), implying that households do not undo the TDF asset allocation by adjusting their holdings of other funds.

These changes in portfolio allocation show up in aggregate statistics for the 401(k) system as a whole. Data provided by the Employee Benefits Research Institute (EBRI) and the Investment Company Institute (ICI) show that between 2006 and 2013 the share of TDFs tripled from 5 percent from 15 percent, while the share of concentrated equity (company stock) declined from 11 percent to 7 percent (a continuation of a longer-term trend). Overall holdings of diversified equity, including equity funds and the equity portion of balanced funds and TDFs but excluding company stock, were little changed over this period (58 percent in both years, with a dip in the period around the financial crisis), but the fraction of plan participants with little or no diversified equity exposure—that is, with less than 20 percent of their portfolios in diversified equity—declined from 26 percent to 18 percent.

TDFs also appear to have altered the pattern of risk-taking by age.\textsuperscript{[Figure 5]} plots the age tilt in several variables, defined as the difference between the average value of the variable for plan participants in their twenties and the average value for participants in their sixties. The share of TDFs develops a positive age tilt in line with the growth of TDF adoption in the population as a whole, since younger participants are more likely to invest in TDFs. During the period of strong TDF growth since 2006, the age tilt in equity holding has increased (with some falling back recently), and the age tilt in low (less than 20 percent) equity participation has declined.\textsuperscript{32} Thus, TDFs have moved overall asset allocation in the direction recommended by standard finance theory. These patterns, and the findings of Mitchell and Utkus (2012), are consistent with the view suggested by Figure 3 that younger households, in particular, are aware of their limited financial knowledge and responsive to the provision of low-cost implicit investment advice.

Despite these favorable results, there are legitimate concerns about the role of TDFs in the post-PPA US retirement system. One is the level of fees. Although TDFs are typically cheaper than the brokers or financial advisers that they often replace, there is meaningful variation in fees for these products just as there is for plain-vanilla index funds.\textsuperscript{33} A related concern is

\textsuperscript{31}This life-cycle pattern is consistent with the recommendations of portfolio choice theory under the assumption that labor income is relatively safe, so that human capital is an implicit holding of a riskless asset that declines as retirement approaches (Campbell and Viceira 2002).

\textsuperscript{32}The large decline in this age tilt in 2008 is driven both by the spread of TDFs among younger participants, and by the response of older participants to the global financial crisis.

\textsuperscript{33}Chalmers and Reuter (2015) compare fees and performance for brokers and TDFs within an Oregon retirement plan. They find average broker fees of 0.90 percent and inferior net performance of broker-advised portfolios.

\textsuperscript{}
the presence of idiosyncratic risk within supposedly diversified TDFs. Balduzzi and Reuter (2015) document considerable heterogeneity in both systematic risk exposures and idiosyncratic returns in TDFs with similar target retirement dates. They attribute this to a tendency for employers to chase performance when they select fund families for their 401(k) plans, and for new entrants in the TDF market to seek market share by taking risk, knowing that they face greater rewards for outperformance than penalties for underperformance.

A deeper unsettled issue is what institutional structure is needed to control conflicts of interest in the provision of retirement savings products, when consumers may not be sufficiently sophisticated to protect themselves. The Department of Labor is currently considering an expansion of the fiduciary standard for retirement investment advisers, with particular relevance for those advising the allocation of individual retirement accounts (IRA). And it is reasonable to ask whether employers, who are often treated as benevolent in behavioral analyses of retirement savings, have incentives and opportunity to exploit behavioral biases of employees when they design retirement benefits. More stringent fiduciary standards are likely to raise costs for rational households, but—as shown by the model of the previous section—they may still be justifiable if the benefits are large enough for unsophisticated households.

**B. Unsecured Consumer Credit**

Consumers with short-term borrowing needs can access several sources of unsecured credit, including credit cards, bank overdrafts, and payday loans. While credit availability provides important budget relief to households facing unexpected emergencies, there are serious questions about consumers’ ability to manage these credit products. These fall into three main categories, each of which maps naturally into one of the types of financial ignorance discussed earlier in this lecture.

First, the structure of fees is sometimes hard to understand. This is a particular concern for credit cards and overdrafts, which offer a complex package of product features. Before the passage of the CARD Act of 2009, credit card issuers could alter interest rates on existing balances and charge substantial fees for late payments or overlimit borrowing. Similarly, banks can charge significant fees for overdraft protection of checks, ATM withdrawals, and debit card transactions. Such fees depend on the account balance at each point in time, which can be hard for consumers to track given bank delays in making deposited funds available and (historically) bank discretion in ordering transactions that reduce balances. As a notorious example, a sequence of many small debits and one large debit will generate much larger overdraft fees if the large debit is processed first, exhausting the account balance, than if the small debits are processed first while the account still has a positive balance.

Second, even if fees are well understood—as is likely the case for payday loans, whose fees are paid on a stand-alone basis—consumers may not appreciate the implications for the cost of credit. Payday loans charge fees of $15 to $20 for $100 loans over two to four weeks, implying annualized percentage rates (APR) of many hundred percent. While the Truth in Lending Act (TILA) of 1968 requires rate disclosure, it is not clear that this is always effective. Consumers with low financial literacy may find it hard to understand the mapping between dollar costs and interest rates.

Third, consumers may not be good at predicting their own future behavior and its implications for the cost of credit. Repeat borrowing by a minority of borrowers generates a large fraction of the total fee revenue for short-term credit products (CFPB 2014). If consumers fail to anticipate their repeat borrowing behavior, they may radically underestimate the cost of an initial loan.

Regulatory responses to these problems have taken a variety of forms. In all cases the challenge is to help consumers who are making mistakes in their credit usage without imposing too great a cost on consumers who are benefiting from the
availability of short-term credit. The simplest strategy is to require disclosure, for example the Schumer box with its standardized credit card terms or the CARD Act’s requirement for credit card statements to compare the total cash payments implied by minimum monthly payments with those implied by paying off credit card debt in 36 months. A second strategy is to ban certain fees unless consumers affirmatively opt in to the associated service. The Federal Reserve required this for overdraft protection of ATM and debit card transactions and the CARD Act required it for credit card overlimit fees, both effective in 2010. A third strategy is to limit the permissible size of fees to a “reasonable” level. The CARD Act has provisions of this sort for credit card late fees, and bans inactivity fees altogether. Some states limit the maximum rate on payday loans through usury laws, and the UK Financial Conduct Authority (FCA) has imposed a cost cap on the UK equivalent of payday lending. A fourth strategy is to restrict the permissible form of credit to limit the extent to which consumers can repeatedly roll over loans, or to encourage an orderly process of paying off debt. For example, in 2010 the state of Colorado required payday loans to be six-month installment loans rather than lump-sum loans (Pew Charitable Trusts 2013; Zingales 2015), and in 2015 the CFPB published a draft proposal that would require payday lenders to limit repeat lending and provide affordable repayment options for rolled-over loans.

An important function of economic research is to measure the effectiveness of such interventions using modern techniques for identification. Researchers have used field experiments to measure the impact of implicit and explicit disclosures on credit usage (Bertrand and Morse 2011; Stango and Zinman 2014). The CARD Act, by regulating personal credit cards but not otherwise similar small-business credit cards, created a natural experiment that is exploited in Agarwal et al. (2015) building on preliminary analysis in CFPB (2013). Agarwal et al. find that the Act considerably reduced the fees paid on credit cards, particularly by borrowers with lower credit scores, without increasing interest rates or reducing the volume of credit. An important unresolved issue is to what extent this reduced bank profits, and to what extent interest charges on higher card balances offset the reduction in fee revenue to banks. Variation in payday lending regulation across states, and within states over time, can be used to identify its effects (Melzer 2011; Carrell and Zinman 2014; Bhutta, Goldin, and Homanoff 2015).

Economic research can also guide policy design, but this requires a deeper understanding of mechanisms. In terms of the stylized theoretical model laid out in the previous section, the key questions are how to measure the fraction of behavioral households who are wrongly choosing or misusing a credit product (b), the size of the damage caused by misuse (normalized to 1 in the model), the benefits of the product when properly used (h), the deadweight cost of intervention (α), and the effectiveness of an intervention that encourages proper usage (θ). Consumer financial regulatory agencies undertake this type of research both to inform regulatory design and, in some cases, to meet the requirements for cost-benefit analysis of specific proposed interventions (Jackson and Rothstein 2015). For example, in order to implement a UK legislative mandate to impose a payday lending cost cap, the FCA used regression discontinuity analysis to compare outcomes for borrowers who fell just below or just above a cutoff level for internal credit scores used by payday lenders. These outcomes were measured using credit bureau files on almost two million first-time loan applicants, and a follow-up survey of 2,000 applicants.

In the United States, the CFPB has studied the problem of repeated overdrafting by a small minority of bank customers (CFPB 2014). The CFPB’s online complaint system can also be understood in these terms as a way to gather data on ex post regret by borrowers, which may be a symptom of behavioral bias in borrowing or of failure to understand fees and borrowing costs. Academic economists can contribute by developing better theoretical models of the interactions between rational consumers, behaviorally

35 See https://www.the-fca.org.uk/price-cap-high-cost-short-term-credit/.
biased consumers, and financial services providers, and by devising credible ways to estimate the parameters of these models.

C. Reverse Mortgages

As a final example of developments in household finance, I briefly discuss the market for reverse mortgages. These are credit products intended to ease the task of funding retirement for households with limited financial assets but substantial housing equity. As Poterba, Venti, and Wise (2011) emphasize, this financial situation is quite common among older US households, and it is likely even more common in countries where housing plays a larger role in household balance sheets as illustrated in Table 2.

The simplest way for a household to tap accumulated housing equity is to sell its home and move to a smaller home or to rental accommodation. However, many older people do not wish to move, and this creates a market for collateralized lending against home equity. This can be done through home equity lines of credit or second mortgages, which are also used by younger borrowers, but reverse mortgages are easier to qualify for and are specifically tailored to the needs of older homeowners. A reverse mortgage has no monthly payments, so the principal balance does not decline over time as it does in a conventional mortgage. Instead, the principal increases at least as fast as the interest rate, and may increase even faster if the loan is structured as a line of credit which the borrower gradually draws upon. The accumulated principal is repaid when the home is sold, and the lender has no recourse in the event that the sales proceeds of the home fail to cover this amount. However, the borrower is responsible for payments of property taxes and homeowner’s insurance, and can be foreclosed upon if these payments are not made.

Almost all reverse mortgages are insured by the Federal Housing Administration (FHA) through the Home Equity Conversion Mortgage (HECM) program. This provides assurance to borrowers that credit will be available in the future under the terms of the reverse mortgage, and it insures lenders against shortfalls caused by negative home equity at the time the house is sold and the mortgage paid off. Given the support to the market provided by FHA insurance, the reverse mortgage market is surprisingly small. A study conducted by the CFPB under the terms of the Dodd-Frank Act (CFPB 2012) found that fewer than 3 percent of eligible households had taken out reverse mortgages in 2011.

Reverse mortgages carry higher interest rates than home equity loans, and this has created a perception that reverse mortgages are expensive. Davidoff (2015) argues that reverse mortgages carry a valuable put option because of their no-recourse feature. The put can be separated from the loan, with its high interest rate, by tapping a line of credit only in circumstances where the value of the home has fallen below the credit available, and where the borrower is willing to default and move. However, Davidoff and Wetzel (2014) find that such “ruthless default” is rare, with many borrowers leaving credit untapped even when additional credit would be effectively free to them. In the absence of an optimal default strategy, reverse mortgages do appear expensive (Lucas 2015) as well as difficult to understand. Both factors limit the demand for the products.

On the other side of the market, lenders appear reluctant to originate reverse mortgages without the benefit of FHA support. This may be because banks cannot be sure that homeowners will not learn to default ruthlessly in the future, even if they fail to do so today. It may also be because banks fear undermaintenance by older homeowners, with negative effects on the quality of collateral (Melzer forthcoming) that can become particularly serious because reverse mortgages offer only limited foreclosure rights to lenders.

Putting these observations together, the reverse mortgage market appears ripe for financial innovation, and there may be a role for public policy to encourage the development of simpler, more affordable products. For example, reverse mortgages might be particularly appropriate vehicles for automatic adjustment of principal balances in response to movements in local house price indexes. Retirees are on a downward trajectory of housing consumption and thus should wish to lay off some house price risk, which is often not the case for younger homeowners (Sinai and Souleles 2005). Adjustment of debt principal to house price indexes preserves a homeowner’s incentives to maintain the house even in a falling housing market. It also eliminates the problematic default option that relatively few homeowners know how to exercise and that borrowers collectively appear to undervalue.
Another design principle might be to combine reverse mortgages with other products attractive to seniors. One reason why retirees retain housing equity may be that they regard it as a vehicle for funding emergency medical expenses and long-term care. A negative health shock overcomes any reluctance to move, and home equity becomes available at the right time to finance spending.[39] This issue could be addressed by marketing reverse mortgages in combination with long-term care insurance and possibly annuities as well, since packaging these products is likely to mitigate the problem of adverse selection.

Lucas (2015) argues that reverse mortgages are expensive in part because of the high costs of reaching older homeowners and educating them about the product. This is an example of the general phenomenon that marketing socially desirable financial products has external benefits to the extent that it establishes favorable social norms and improves financial literacy. In this circumstance, consumer financial regulation may be able to play the positive role of promoting desirable innovation rather than the negative role of restraining undesirable innovation.

V. Conclusion

The complexity of twenty-first century financial arrangements poses a daunting challenge to households managing their financial affairs, to regulators attempting to assist them, and to the economics profession. In this lecture I have argued that household financial mistakes create a new rationale for intervention in the economy. I have emphasized that beyond the easy cases where behaviorally biased households can be “nudged” to avoid mistakes with minimal effects on rational households, financial regulators face a difficult trade-off between the benefits of regulation to households that make mistakes, and the costs of regulation to other financial market participants. The task for economists is to confront this trade-off explicitly, bringing to bear the highest quality evidence that modern applied microeconomics can make available.

REFERENCES


39 Consistent with this, Nakajima and Telyukova (2015) show that retirees run down both financial and housing wealth faster in countries with greater public funding of health care, and Ameriks et al. (2015) find that Vanguard clients report health care concerns to be a major reason for their limited dissaving.


Calvet, Laurent E., and Paolo Sodini. 2014. “Twin Picks: Disentangling the Determinants of


Johnson, Eric, Stephan Meier, and Olivier Toubia.


Poterba, James, Steven Venti, and David Wise.
MAY 2016

30

AEA PAPERS AND PROCEEDINGS

MAY 2016

30

AEA PAPERS AND PROCEEDINGS


