Visibility Bias in the Transmission of Consumption Beliefs and Undersaving[*]

Bing Han[†] David Hirshleifer[‡] Johan Walden[§]

October 30, 2020

Abstract

We model visibility bias in the social transmission of consumption behavior. When consumption is more salient than non-consumption, people perceive that others are consuming heavily, and infer that future prospects are favorable. This increases aggregate consumption in a positive feedback loop. A distinctive implication is that disclosure policy interventions can ameliorate undersaving. In contrast with wealth-signaling models, information asymmetry about wealth *reduces* overconsumption. The model predicts that saving is influenced by social connectedness, observation biases, and demographic structure; and provides a novel explanation for the dramatic drop in savings rates in the US and several other countries in recent decades.

JEL Classification: G11, G41

[*]We thank Pierluigi Balduzzi, Martin Cherkes, Marco Cipriani, Douglas Bernheim, Andreas Fuster, Giacomo de Giorgi, Mariassunta Giannetti, Luigi Guiso, Robert Korajczyk, Theresa Kuchler, Chris Malloy, Timothy McQuade, Peter Reiss, Nick Roussanov, Erling Skancke, Andy Skrzypacz, Tano Santos, Siew Hong Teoh, Bernie Yeung, Youqing Zhou, David Benjamin Zuckerman, and participants at the Amerian Economics Association meetings in Boston, the MIT Conference on Evolution and Financial Markets, the National Bureau of Economic Research Behavioral Finance Working Group meeting at UCSD, the OSU Alumni Conference at Ohio State University, the Miami Behavioral Finance Conference at the University of Miami, the SFS Finance Cavalcade at the University of Toronto; the WISE18 Workshop on Information and Social Economics at Caltech; and seminar participants at Arizona State University, Boston College, Calgary University, Caltech, Cheung Kong Graduate School of Business, Columbia University, Hong Kong Polytechnic University, London Business School, National University of Singapore, New York Federal Reserve Board, Peking University, Renmin University, Shanghai University of Finance and Economics, Southwest University of Finance and Economics, Stanford University, University of Hong Kong, University of Houston, University of International Business and Economics, University of Texas Austin, and York University for very helpful comments; and Lin Sun and Qiguang Wang for very helpful comments and research assistance.

[†]Rotman School of Management, University of Toronto.

[‡]Merage School of Business, UC Irvine, 4291 Pereira Drive, SB2 Suite 406, Irvine, CA 92697. email:david.h@uci.edu.

[§]Haas School of Business, UC Berkeley.

1 Introduction

Several authors have argued that people have little idea how much they should save for retirement (e.g., Akerlof and Shiller (2009)), owing either to lack of relevant information, or failure to process it effectively. It is hard to know what stream of satisfaction will actually result from a consumption/savings rule chosen today.[1] Future wealth realizations are risky, and it is hard to forecast remaining lifespan or health in old age. For the latter, most people do not process the relevant public but technical information contained in mortality tables and medical research.

This suggests that people are often 'grasping at straws' in their savings decisions, and that people look to social cues for help.[2] (We discuss evidence about saving and psychology that motivate our model in Section 2). Surprisingly, however, there has been little formal modeling of how biases in social learning processes affect lifetime consumption/savings choices. The nature of such effects is not immediately obvious. There is evidence of contagion of consumption and investment behaviors, but contagion can potentially spread either a decision to consume more or a decision to consume less. Notably, little is known about whether biased social learning implies over- versus underconsumption, and whether policy interventions can help remedy such a *directional* bias.

We address this topic in a model in which people are more likely to observe potential consumption events that do rather than do not occur. For example, a boat parked in a driveway draws the attention of neighbors more than the absence of a boat. Similarly, it is more noticeable when a friend or acquaintance is encountered eating out or reports taking an expensive trip than when not, or buys an enjoyable product as compared with not doing so. We call the greater availability and salience of engaging in a consumption activity *visibility bias.*

We further assume that people do not adequately adjust for the selection bias toward noticing the consumption rather than nonconsumption events of others. This causes undue updating toward the belief that others are consuming heavily. So observers conclude that future consumption prospects are good, and therefore that low saving is appropriate. Observers therefore choose a high level for their own actual consumption. We refer to average levels of consumption higher

[1] Allen and Carroll (2001) point out that "...the consumer cannot directly perceive the value function associated with a given consumption rule, but instead must evaluate the consumption rule by living with it for long enough to get a good idea of its performance. ... it takes a very large amount of experience ... to get an accurate sense of how good or bad that rule is.'

[2] A large analytical literature shows that socially inefficient outcomes can arise from rational or biased social learning (e.g., see the survey of Golub and Sadler (2016)), and there are models of how social interaction affects investment and saving behaviors such as market participation, house purchases, and the aggressiveness of trading in individual stocks (Hong, Kubik, and Stein 2004; Burnside, Eichenbaum, and Rebelo 2016; Han, Hirshleifer, and Walden 2020).

than would occur with zero visibility bias as *overconsumption*.

As a result, visibility bias effects are self-reinforcing. Each agent becomes an overconsuming model for others. This positive feedback can cause severe undersaving in society as a whole, even when visibility bias is mild. Also, in market equilibrium, the reluctance of agents to save results in a higher interest rate.

Furthermore, in a social network setting, we find that such effects are amplified when in-degree is positively associated with out-degree, as is likely the case since people differ in their overall sociability. Intuitively, agents who observe others heavily (high in-degree) will be especially influenced by visibility bias, causing them to overconsume heavily. If such agents are the ones who are most heavily observed by others (high out-degree), then the heavy overconsumption of these agents is especially influential for others as well.

The two premises of our model—that consumption activities are visibility biased (more available and salient to others than nonconsumption); and that people do not adequately adjust for selection bias in their attention toward consumption—are motivated by the psychology of attention, salience, and social communication (see Section 2). Visibility bias in our model need not be viewed as a cognitive failure; it is a source of bias in the *social transmission* of information. There are good reasons to allocate more attention to occurrences (more generally, to salient events) than to nonoccurrences. However, failing to adjust appropriately for this selection bias in attention/observation *is* an error—one that produces a directional bias in inferences.

The model also has implications for how saving rates change through time. A well-known puzzle is that personal saving rates in the U.S. have declined dramatically since the 1980s, from 10% in the early 1980s to a low of about 3% in 2007, while national debt has increased. This has raised concerns among many observers about whether Americans will be able to sustain their standards of living in retirement. A similar trend has occurred in many OECD countries, with ratios of household debt to disposable income often reaching well over 100% (OECD 2014).

The visibility bias approach offers a novel explanation. The model is driven by observation of the consumption of others; greater observability of consumption intensifies the overconsumption effect. The rise of electronic communications reduced the cost of observing the behaviors of distant individuals. For example, the drop in costs of cell phones and long-distance calls, the rise of cable television and VCRs (video cassette recorders), and subsequently the rise of the internet, greatly increased people's ability to observe others' consumptions, as people are able to hear, view, or report via social networks about consumption experiences.[3] So even prior to the internet, the rise

[3]The rise of an increased diversity of cable television offerings (including channels devoted to shopping, travel,

of electronic communications was transforming social observation of others' consumption. Such biased observation is the driving force behind overconsumption in our model.

There are also notable differences in savings rates across countries and ethnic groups which are not well-explained by traditional economic models (Bosworth 1993). A new possible explanation suggested by our approach is that cultural differences affect communication about, or observability of others' consumption or wealth.[4] Our model also implies that urbanization will be negatively related to saving, as urbanization is associated with a higher intensity of social interaction and observation of the consumption of others. This prediction (which derives from the *social* feature of our approach) is consistent with the evidence of Loayza, Schmidt-Hebbel, and Serven (2000).

Overconsumption in our approach derives from underestimation of the risk of adverse economic shocks. Under some assumptions, there are also biased beliefs about the behaviors of others. Such mistakes can potentially be corrected. So a distinctive empirical and policy implication of the visibility bias approach is that salient public disclosure of accurate information about wealth risks, such as of layoffs or of high health care bills, can help reduce overconsumption.

However, in practice, announcements of probability estimates may be hard for people to process and convert into consumption plans. This suggests saliently disclosing information about *how much others actually consume*. Under appropriate conditions, such disclosures reduce overconsumption. Furthermore, accurate disclosures that make saving behavior more salient (i.e., disclosures that are visibility-biased toward saving) can be especially effective in reducing overconsumption.

The model offers new insight into extensive evidence from social psychology and economics that people often have biased perceptions about the popularity of different attitudes and behaviors of others. Social psychologists have argued that overestimation of the popularity of a behavior causes people to wrongly perceive it to be normative, and that in some cases disclosure interventions can help remedy the problem.

In our model, agents neglect the selection bias toward observing consumption events, and therefore update toward overoptimistic beliefs about future prospects that favor higher consumption. Nevertheless, the equilibrium consumption levels confirm agents' high beliefs about others' consumption. So there is, on average, no overestimation of others' consumption. The model there-

---

home remodeling, and other costly leisure pursuits, as well as dramas that indirectly highlight consumption activities) further increased visibility. People often report by phone or other electronic networks on such activities as traveling, eating out, and recent product purchases. Social media for sharing pictures and videos, such as Instagram, have heavy emphasis on travel, fashion, and celebrities, all of which are associated with high observation of others' consumption.

[4] Carroll, Rhee and Rhee (1994, 1999) provide divergent evidence as to whether culture affects savings. Guin (2017) provides evidence of a strong effect of culture, proxied by language, on saving.

fore provides a surprising contrast with the intuition from the abovementioned social psychology literature, by showing how mistaken learning from others can increase a behavior without any overestimation of how much others engage in it.

In the special case of the base model, in which all agents are ex ante identical, disclosure of others' average consumption does not correct agents' average beliefs. However, when there are heterogeneous agents, we explore the conditions under which different types of agents either overestimate or underestimate others' consumption. These provide insight into the conditions under which disclosure increases saving, and about how to attune the disclosure policy to these conditions to maximize effectiveness. For example, in reality some people have more accurate prior knowledge than others about the risk of adverse wealth shocks, or are less naive than others about visibility bias. In a simple special case of the model with such "smart agents," a disclosure of actual average consumption causes non-smart agents on average to revise their consumptions downward more than smart agents revise upward. This reduces average consumption. The effect of disclosure on consumption is driven by differences across types in their prior knowledge. There is evidence consistent with this direction of effect (D'Acunto, Rossi, and Weber (2020)).

Furthermore, in reality some people observe social signals more heavily than do others (relative to information not obtained by observing others' consumption). Holding all else equal, we find that agents who observe others heavily overconsume more than those who observe others lightly. In consequence, public disclosure of the low average consumption *of those who observe others lightly* on average pulls down the consumption of heavy observers, reducing per capita consumption. But in surprising contrast, disclosure of *population-wide* per capita consumption *increases* overconsumption. This is because a greater number of social observations gives an agent higher subjective certainty about the risk of adverse wealth shocks. So the disclosure of aggregate consumption pulls upward the consumption of the light observers more than it pulls down the consumption of heavy observers. This effect is driven by differences across types in their amount of social observation.

We call agent types that on average consume less than other types at date 0 low-consumers, and agent types that on average consume more than other types at date 0 high-consumers. Then overall, a key determinant of whether disclosure of aggregate consumption reduces average consumption is whether high consumers on average have higher or lower subjective confidence about possible wealth shocks than low consumers. If a type is low-consuming because this type has greater genuine knowledge (relatively high observation of many unbiased private signals), the low consumers will have higher confidence in their beliefs. But if a type is low-consuming because this type makes relatively few social observations, the low consumers will have lower confidence in their beliefs.

So empirically, the predicted effect of disclosure of aggregate consumption depends on the dispersion across agents in amount of unbiased information they possess, versus dispersion in how much they observe others. If agents differ mostly in amount of unbiased information (number of nonsocial signals), then we expect disclosure to reduce per capita consumption. If instead agents differ mostly in how heavily they observe others, then we expect disclosure to increase per capita consumption. Empirically, there are proxies for both the amount of social observation engaged in by different individuals (e.g., using social media or survey data), and individuals' degree of education or quantitative sophistication. So these predictions are potentially testable.

The conclusion that agents overconsume when they are young, at the expense of consumption when old, extends to an overlapping generations setting in which the young can observe old as well as young agents. Overconsumption by the young is decreasing with the extent to which their observations are tilted toward the old. This tilt depends on the age distribution of the population, and on how visible and salient consumption by members of the two groups is to young observers. These distinctive empirical implications of our approach are as yet untested. Furthermore, since the old on average consume less than the young, in this setting there is another policy intervention that can reduce overconsumption—salient disclosure of the consumption of the old.

We are not the first to study how psychological bias affects consumption. A plausible alternative theory of overconsumption and undersaving is that people are present-biased (i.e., subject to hyperbolic discounting, Laibson (1997)). Present bias is a preference effect, whereas the visibility bias approach is based on belief updating. Also, present bias is an individual-level bias, whereas the visibility bias approach is based upon social observation and influence. The visibility bias approach therefore has the distinctive implications that the intensity of social interactions and shifts in the technology for observing the consumption of others affect how heavily people consume. It also implies that population level characteristics such as population density, sociability, demographic structure, and wealth dispersion affect average consumption outcomes, in contrast with approaches based upon pure individual-level biases.

Another appealing approach to overconsumption is based on Veblen effects (Cole, Mailath, and Postlewaite 1995; Bagwell and Bernheim 1996; Corneo and Jeanne 1997; Charles, Hurst, and Roussanov 2009), wherein people overconsume to signal high wealth to others. In such models, beliefs are rational, whereas the visibility bias approach is based upon biased updating. The visibility bias approach has distinct empirical implications as well. Information asymmetry about others' wealths is the source of Veblen effects, which are not present when wealths are equal. In contrast, as shown in Section 4.2, in the visibility bias approach, overconsumption is strongest when there is *low* wealth dispersion and information asymmetry about wealth.

A third approach is based on agents deriving utility as a function of the consumptions of other agents (Abel 1990; Galí 1994; Campbell and Cochrane 1999), or alternatively having payoff complementarities between the actions of different agents. The concern for relative consumption is often referred to as the 'keeping up with the Joneses' approach. These approaches do not in general necessarily imply overconsumption (Dupor and Liu (1993), Beshears et al. (2018)), but this does arise in some settings (Harbaugh 1996; Ljungqvist and Uhlig 2000). In addition to unambiguously predicting a specific direction of effect, overconsumption, the visibility bias approach offers various distinctive implications about the effects of disclosure and of shifts in visibility bias, wealth dispersion, and demographics.

We have mentioned that a further empirical and policy implication of the visibility bias approach is that salient public disclosure can help correct people's beliefs, reducing overconsumption. That a relatively simple policy intervention can potentially ameliorate the undersaving problem is a distinctive feature of the visibility bias approach to the consumption/saving decision.

Jackson (2019) considers a setting with utility or payoff interactions that generate strategic complementarity between agents' actions, a possible example being recreational drug use as a social activity. In Jackson's model, misestimation of the average actions of others affects behavior, which can also occur in our setting. However, in our setting there is no strategic complementarity—utility of consumption does not depend upon others' consumption. So we provide a model of general consumption and savings levels, rather than of those activities that have positive strategic complementarities. Also, Jackson's model is based on the interplay between strategic complementarity and the "friendship paradox" in social networks. In contrast, our main results (e.g., overconsumption) do not rely on the friendship paradox.

Finally, another approach that can lead to overconsumption is based on speculative disagreement; see Heyerdahl-Larsen and Walden (2017). When investors with heterogeneous beliefs bet against each other in an asset market, they may all expect to profit, at least some of them mistakenly. Depending on agents' elasticity of intertemporal substitution, this can result in equilibrium overconsumption. Several of the implications discussed above also distinguish our approach from theirs. For example, the speculative disagreement approach does not share the implications here about network properties and overconsumption.

2 Motivating Evidence

We next discuss evidence that motivates our approach, starting with evidence about saving behavior. In deciding how much to save, people make very basic mistakes, and rely on noisy cues.[5] There is also considerable evidence that social interactions affect several dimensions of consumption, saving, and investment choices.[6]

Based on evidence about retirement investment, many economists argue that US households undersave, though there are other viewpoints.[7] A further notable stylized fact in household finance is the drop in the savings rate in the U.S. and several other countries over a period of decades.[8] Our model offers an explanation for both phenomena.

We next turn to institutional background and evidence from the psychology of attention and salience that motivate our modeling approach. The two key assumptions of our model are that consumption activities are more available and salient to others than nonconsumption; and that people do not adequately adjust for the selection bias in their attention toward these consumption events.

With regard to the first assumption, there is extensive evidence that occurrences are more salient and more fully processed than nonoccurrences (e.g., Neisser (1963), Healy (1981), the review of Hearst (1991), and Enke (2020)). Occurrences provide sensory or cognitive cues that trigger attention. In the absence of such triggers, an individual will only react if (as is usually not the case) the individual is actively monitoring for a possible absence. This is what is striking about the famous phrase "The dog that did not bark" in the Sherlock Holmes story; his stroke of genius is to recognize the importance of an absence. An example of the low salience of non-occurrences is neglect of opportunity costs, i.e., hypothetical benefits that would occur under alternative courses of action. Consistent with the application of these ideas to consumption, Frederick (2012)

[5] See Samuelson and Zeckhauser (1988), Shefrin and Thaler (1988), Madrian and Shea (2001), Beshears et al. (2008), Benhassine et al. (2015).

[6] Duflo and Saez (2002, 2003), Hong, Kubik, and Stein (2004), Brown et al. (2008), Charles, Hurst, and Roussanov (2009), Kaustia and Knüpfer (2012), Georgarakos, Haliassos, and Pasini (2014), Shemesh and Zapatero (2016), Giorgi, Frederiksen, and Pistaferri (2016), and the evidence reviewed in Hirshleifer and Teoh (2009).

[7] See Laibson, Repetto, and Tobacman (1998), Madrian and Shea (2001), Poterba, Venti, and Wise (2012), Poterba (2014), U.S. Government Accountability Office (2015), Stanford Center on Longevity (2016), and Gomes et al. (2020), and the divergent viewpoint of Scholz, Seshadri, and Khitatrakun (2006)). Over 50% of U.S. adults say that they could not easily come up with $400 to cover an emergency expense (Federal Reserve Board 2018). More than half of households with bank cards carry debt from month to month, almost always at high interest rates; a substantial fraction borrow at close to their credit limits (Gross and Souleles 2002).

[8] Some review articles argue that this drop is hard to explain with existing models (Parker 1999; Guidolin and Jeunesse 2007), but economists have proposed a wide range of potential explanations. Parker (1999) concludes that "Each of the major current theories of the decline in the U.S. saving rate fails on its own to match significant aspects of the macroeconomic or household data." Guidolin and Jeunesse (2007) argue that factors such as greater capital mobility, new financial instruments, and aging populations do not suffice to explain the phenomenon, and conclude: "The recent decline of the U.S. private saving rate remains a puzzle."

concludes that "purchasing and consumption are more conspicuous than forbearance and thrift," and gives the example that "Customers in the queue at Starbucks are more visible than those hidden away in their offices unwilling to spend $4 on coffee."

One reason that consumption activities are highly visible is that many are social, such as eating at restaurants, wearing stylish clothing to work or parties, and traveling. Furthermore, physical shopping is itself a social activity. Shopping and product evaluation are also engaging topics of conversation. Many television dramas display glamorous consumption activities, travel, entertaining, and dining, and some media channels explicitly focus on shopping and other costly leisure activities. In contrast, saving for retirement is typically a private activity with very low visibility to others.

With regard to the second key assumption of our model, evidence from both psychology, experimental economics, and field studies of *selection neglect* confirms that observers often fail to adjust appropriately for data selection biases (Nisbett and Ross 1980; Brenner, Koehler, and Tversky 1996).[9] Neglect of absences is also reflected in the principle of WYSIATI, "What you see is all there is," one of the key features of System 1 thinking (Kahneman 2011). In general, neglect of selection bias is implied by the representativeness heuristic of Kahneman and Tversky (1972). Owing to limited cognitive resources, adjusting for selection bias requires attention, and effort. Selection bias is especially hard for people to correct for because adjustment requires attending to the non-occurrences that shape a sample. The combination of visibility bias and selection neglect in our model can be viewed as endogenizing the *availability heuristic* of Kahneman and Tversky (1973), so the tendency in the model to update toward thinking others are consuming heavily can alternatively be interpreted as coming from the use of this heuristic.[10]

Potentially consistent with the idea that the combination of visibility bias and selection neglect distort perceptions, Frederick (2012) provides experimental evidence that the salience of consumption results in overestimation by observers of how much other individuals value certain consumer products. Consistent more broadly with the idea that visibility bias affects consumption behavior, there is evidence that people are influenced in car purchase decisions by observation of the purchases of others (Grinblatt, Keloharju, and Ikäheimo (2008), Shemesh and Zapatero

[9] People often naively accept sample data at face value (Fiedler 2008). Mutual fund families advertise their better-performing funds; in the experimental laboratory both novice investors and financial professionals misinterpret reported fund performance owing to selection neglect (Koehler and Mercer 2009). Auction bidders in economic experiments tend to suffer from the winner's curse (neglect of the selection bias inherent in winning), and hence tend to lose money on average (Parlour, Prasnikar, and Rajan 2007).

[10] According to the availability heuristic, people overestimate the frequency of events that come to mind more easily, such as events that are highly memorable and salient. The availability heuristic is therefore a failure to adjust for the selection bias in information brought to conscious attention—this being the subset of information that was stored into memory and is easy to retrieve from it (e.g., consumption rather than nonconsumption activities).

(2016)), and that effects are stronger in areas where commuting patterns make the cars driven by others more visible (McShane, Bradlow, and Berger 2012). There is also evidence that the observability of others' choices is important for social influence in financial decision-making (Lieber and Skimmyhorn 2018).

Generally it is more interesting to hear about an action than inaction. So those who consume may chooses to discuss their action more than those who do not. As a result, the consumption activities of others may be more cognitively available than non-consumption. Berger and Milkman (2012) provide evidence that online content is more likely to go viral when it is positive than negative, and more rather than less arousing. This evidence suggests that people are more prone to sharing news about consumption activities, which are enjoyable and arousing, than news about stoical restraint from consuming. Finally, survey evidence is consistent with several ingredients of the model's line of reasoning for why there is overconsumption as outlined in the introduction.[11]

3 The Model

To capture some basic insights parsimoniously, we start with our primary framework, and consider extensions in Section 4. Consider an economy with N agents, where N is large. Each agent maximizes a quadratic expected utility function with zero subjective rate of discount over two dates,

$$U = c_0 - \left(\frac{\rho}{2}\right) c_0^2 + E\left[c_1 - \left(\frac{\rho}{2}\right) c_1^2\right], \tag{1}$$

where c_0 and c_1 are consumptions at dates 0 and 1. We permit possible negative consumption, $c_1 < 0$. The two dates can be viewed as reflecting consumption early versus late in the life cycle.

At date 0, each agent chooses how much to consume and how much to borrow or lend at the riskfree interest rate $r = 0$, so each agent's budget constraint is

$$c_1 = W - c_0 - \epsilon, \tag{2}$$

[11] Consistent with high salience of consumption activities, the 2019 Modern Wealth Index Survey by Charles Schwab finds that three in five Americans pay more attention to their friends' spending activities than friends' saving. Consistent with observation of others affecting behavior, nearly half of millennials (49%) say that their spending habits have been influenced by the photos and experiences their friends share on social media. Consistent with observation of others potentially affecting welfare outcomes, in the Fidelity Investments 2018 Millennial Money Study, 63 percent of social media users report that social media has a negative influence on their financial well-being. Consistent with Fidelity regarding visibility bias as a problem, Fidelity offers the following advice: "focus on your own opportunities, and not on those in your network. Often, the life one portrays on social media does not show the full picture, so take all those photos, snaps, stories and tweets for what they are: a curated snapshot of one moment. Remind yourself to remain focused on your goals, not the moments others may be displaying through a rose-colored filter."

where W is date 0 wealth, and ϵ is a potential wealth shock at time 1. We assume that $\rho W < 1$, to ensure that utility is increasing in consumption. We assume that

$$\epsilon = \begin{cases} 0 & \text{with probability } \; p \\ W & \text{with probability } \; 1-p, \end{cases} \tag{3}$$

so with probability $0 < p < 1$, an agent's date 1 wealth is high, and with probability $1-p$ it is low, where p is common to all agents. There is indeed evidence that fear of adverse wealth shocks strongly affects consumption/savings decisions (Malmendier and Shen 2018).

The distribution of ϵ is the same for all agents, and can have any correlation across agents. So this negative wealth shock can represent a systematic outcome such as a major depression, or an underfunded pension system; or an idiosyncratic outcome that all agents are symmetrically exposed to, such as the possibility of a financially costly illness, disability or job loss. The key is that agents draw inferences about the common probability p of such an event from their observations of the consumption of others.[12]

We will refer to beliefs formed before any social observation as "prior beliefs." Based on prior beliefs and observation of others, each agent forms a probability estimate that date 1 wealth will be high ($\epsilon = 0$), which we denote $\hat{p}$. Based on this estimate, an agent chooses date 0 consumption to maximize expected utility, which by (1) yields optimal consumption

$$c_0 = \hat{p}\left(\frac{W}{2}\right). \tag{4}$$

So date 0 consumption is proportional to the estimated probability that future consumption will be high. If people were sure of the good outcome ($\hat{p} = 1$), they would consume half their total wealth; otherwise they consume less than half.

To obtain implications for how people save out of available income, we can think of wealth, W, as being a discounted value of income that is generated over time. Since the interest rate is zero, the agent's opportunity set as given in equations (2) and (3) is consistent with the agent receiving a cash flow of $W/2$ in each of the two periods, where if the wealth disaster strikes, a further incremental cash flow of $\epsilon = -W$ is obtained at date 1. Under this assumption, the agent's

[12]The possibility of wealth disasters is a convenient way to capture the idea that there is a random outcome that affects the benefits from deferring consumption. Alternative modeling approaches that would yield similar results would have agents learning from others about the probability of dying young, of experiencing adverse health events, or of experiencing rapid salary growth, each of which would also affect the benefits of saving.

saving (at time 0) is

$$s_0 = (1 - \hat{p}) \left(\frac{W}{2} \right). \tag{5}$$

The total date-0 potential consumption of an agent, $W/2$, is divided into K different activities which we call *bins* (K large), where each bin represents potential consumption of $W/(2K)$. We refer to a bin as *full* if it contains consumption and *empty* otherwise. An agent who chose to consume $W/2$ at date 0 (which, by equation (4), is consistent with belief $\hat{p} = 1$, i.e., no risk of a negative shock) would then have all bins full, whereas an agent who chooses to consume 0 (consistent with belief $\hat{p} = 0$, i.e., certainty of the adverse outcome) would have all bins empty. As seen above, neither c_0 nor s_0 depend on the correlation of ϵ shocks across investors, since those shocks are not realized until date 1. So we do not need to make any assumption about the value of this correlation.

There are $G \geq 1$ different *groups*, or *types*, of agents, where group/type can refer, for example, to how well informed an agent is, how heavily an agent engages in social observation, or how heavily the agent is observed by others. For readers who seek a quick basic intuition, the case $G = 1$ is covered in Section 3.1.

The number of agents of each type is large. The fraction of type g in the population is f_g, $g = 1, \ldots, G$, $\sum_g f_g = 1$. Agents have two sources of information that they use to estimate p. First, each agent of type g observes L_g i.i.d. unbiased private signals about p. So different types may observe different numbers of signals. Specifically, agent n observes the private signals $\tilde{x}_k^n$, $k = 1, \ldots, L_g$, where $\tilde{x}_k^n \sim \text{Ber}(p)$ is Bernoulli distributed with expected value p. These signals represent the information an agent gathers through sources other than observing the consumption and savings of others.

Second, each agent observes a subset of other agents' consumption bins. Specifically, an agent n of type g observes M_g independently drawn Bernoulli distributed consumption bins of other agents, $\tilde{y}_k^n$, $k = 1, \ldots, M_g$.

Crucially, we assume that each agent believes that the probability of observing a full bin (indicating that $\tilde{y}_k^n = 1$) is p, i.e., that $\tilde{y}_k^n \sim \text{Ber}(p)$. We will see that when observations of others' bins are unbiased, in equilibrium this belief is correct. So our assumption is consistent with agents believing that all agents are sampling others' consumption bins without bias.

Each agent is Bayesian, with an improper $\text{Beta}(0, 0)$ distributed prior belief about the probability p of a wealth disaster. The Beta distribution is standard for tractably describing how a Bayesian updates beliefs about a Bernoulli probability based on observations of Bernoulli outcomes. From an agent's perspective, each of the observed $L_g + M_g$ signals is equally informative about p. The L_g signals are probability-p Bernoulli signal drawings, and the M_g signals are Bernoulli observations of whether other agents' bins are full, with probability also perceived to be p.

By a standard formula for Bayesian updating of beliefs about a Bernoulli distribution parameter based on observation of Bernoulli outcomes (DeGroot and Schervish (2012), Theorem 7.3.1, p. 395), agent n updates to the posterior belief

$$\hat{p}^n = \frac{L_g \bar{x}^n + M_g \bar{y}^n}{L_g + M_g} \equiv \frac{\bar{x}^n + m_g \bar{y}^n}{1 + m_g}, \tag{6}$$

where $\bar{x}^n \stackrel{\text{def}}{=} \frac{1}{L_g}\sum_{k=1}^{L_g} \tilde{x}_k^n$ is the average of the agent's private signals, and $\bar{y}^n \stackrel{\text{def}}{=} \frac{1}{M_g}\sum_{k=1}^{M_g} \tilde{y}_k^n$ is the average of the agent's social observations. The parameter $m_g \stackrel{\text{def}}{=} M_g/L_g$ is the total weight an agent of type g puts on the consumption bin observations relative to the agent's private signals. Without loss of generality, we order the types so that m_g is increasing in g, $m_1 < m_2 < \cdots < m_G$, i.e., the higher the type, the higher is the weight the type places upon social information.

For most of the paper (with the exception of Section 3.3 on disclosure interventions), we assume that $L_g = L$ for all g, so that the types differ only in the number of signals M_g about the consumption of others that agents observe. In particular, suppose that all of the M_g bins that an agent observes belong to different agents. Then the agent is observing M_g different agents. So using network terminology, m_g is proportional to the *in-degree* of agents of type g. This parameter is a driver of some of the model's distinctive empirical implications.

The bins an agent observes are randomly chosen from the rest of the population with possible overweighting of some types relative to others.[13] The probability that a chosen bin comes from an agent of type g is u_g, $g = 1, \ldots, G$, $\sum_g u_g = 1$. If $u_g = f_g$, this probability matches the population fraction of the type, so that selection is unbiased across types. Otherwise, some types are disproportionately influential.

A key assumption of the model is that a full consumption bin is disproportionately likely to be observed relative to an empty bin, where observation of a full versus empty bin is independent of which agent type the bin is selected from. The higher likelihood of drawing a full bin derives from what we call *visibility bias*, the tendency to notice and recall occurrences rather than non-occurrences.[14] Agents mistakenly form beliefs as if there were no visibility bias. The failure of

[13] Technically, since we assume that the observer never draws bins from the same agent twice, drawings are non-independent. However, since the number of agents of each type is large, this dependence vanishes.

[14] In reality, the occurrence versus non-occurrence distinction that we focus upon is not the only source of differ-

the agent to adjust for this overrepresentation of full bins in their samples is a type of selection neglect. This neglect causes them to update their beliefs incorrectly.

If B^F of the B bins in the population are full and B^E are empty, then the chance that an *observed* bin is full is

$$\frac{k^F B^F}{k^F B^F + k^E B^E} = \frac{\frac{B^F}{B}}{\frac{B^F}{B} + \frac{k^F}{k^E}\left(1 - \frac{B^F}{B}\right)} = \frac{b}{b + \frac{1-b}{\tau}} \stackrel{\text{def}}{=} S_\tau(b), \tag{7}$$

where k^F is the probability that a bin is observed conditional upon it being full, k^E is the probability that a bin is observed conditional upon it being empty, $\tau \stackrel{\text{def}}{=} k^F/k^E \geq 1$, and $b \stackrel{\text{def}}{=} B^F/B$ is the consumption fraction. We call $S_\tau(b)$ the visibility bias function. It plays a central role in our analysis.

The parameter τ measures the overrepresentation of full bins in an observer's sample, i.e., *visibility bias*, where $\tau = 1$ indicates no visibility bias. When $\tau > 1$, there is overrepresentation of draws of consumption bins over non-consumption bins.[15] It follows immediately from (7) that the function $S_\tau(b)$ is strictly increasing in τ and $b \in (0, 1)$, is concave in b under our assumption that $\tau > 1$, and satisfies $S_\tau(0) = 0$, $S_\tau(1) = 1$, and $S(b) > b$ when $0 < b < 1$. In contrast, $S_1(b) \equiv b$.

Letting the average fraction of full bins among agents of type g be denoted $\bar{p}_g$, and recalling that the probability that an observed bin drawn from a type-g agent is u_g, the probability that a random bin of an *observed* agent in the population is full is $\bar{r} = \sum_g u_g \bar{p}_g$, and the probability that a bin *observation* from such an agent is full is $S_\tau(\bar{r})$. If many agents of type g independently observe such draws and update according to (6), by the law of large numbers, their average probability estimate will in the limit almost surely be their expected probability of observing a full bin, which is

$$E[\hat{p}^n] = \frac{p + m_g S_\tau(\bar{r})}{1 + m_g}. \tag{8}$$

Since an agent's consumption in (4) is proportional to the agent's probability estimate, this quantity is proportional (with factor $W/2$) to per capita consumption in the population.

Neglect of visibility bias tends to increase probability estimates, and thereby consumption,

ences in the salience of different consumption behaviors. For example, extreme outcomes also tend to be psychologically salient. Other things equal, we might expect this to cause observers to notice especially when others have either unusually low or unusually high total consumption, with no clear overall bias toward either over- or underestimation of others' consumption. Our modelling focus is on an attentional bias—neglect of nonoccurrences—that has a clear-cut directional implication.

[15]We refer to agents as observing a biased sample of target activities. However, the algebra of the updating process can equally be interpreted as reflecting a setting in which observers draw unbiased random samples of full versus empty bins, but are biased in their ability to *retrieve* different observations for cognitive processing and belief formation.

above that in a rational setting with no visibility bias. With visibility bias, if (out of equilibrium) the average probability estimate of each type were correct, $\bar{p}_g = p$, where $0 < p < 1$, then we would have $r = p$, so by (8) and since $S_\tau(x) > x$, the expected probability estimate of an observing agent of type g would be

$$\frac{p + m_g S_\tau(p)}{1 + m_g} > p, \qquad \text{when } \tau > 1, 0 < p < 1, \tag{9}$$

which would contradict the premise.

An equilibrium is now defined by a set of self-confirming average probability estimates of the different types, $\bar{p}_g$, $g = 1, \ldots, G$, such that

$$\bar{p}_g = \frac{p + m_g S_\tau(\bar{r})}{1 + m_g} \tag{10}$$

$$\bar{r} = \sum_g u_g \bar{p}_g. \tag{11}$$

In other words, for each type, the average probability estimate for agents of that type is the naive Bayesian update based upon observing a biased sample of bins from a population of agents who choose consumption based upon the same specified probabilities $\bar{p}_g$ for each type. The population-average probability estimates and consumption in equilibrium are then

$$\bar{p} = \sum_g f_g \bar{p}_g, \tag{12}$$

$$\bar{c}_0 = \frac{\bar{p}W}{2}. \tag{13}$$

The above argument relies on the law of large numbers, so the number of agents of each type needs to be large. In the appendix, we define equilibrium rigorously as a limiting concept as the economy becomes large. Also, we have presented here a static equilibrium concept in which agents simultaneously choose consumption and observe samples of others' consumption. In the appendix we allow agents to observe each other and make decisions sequentially. We introduce there a growing sequence of economies and study the large economy limit as the number of agents tends to infinity. We divide the initial time period into many sub-periods of very short length. In each period a fraction of agents are selected who observe bins from the consumption of previous sub-period agents and choose their consumption. We show that the subperiod average probability estimate and consumption in the limit converge almost surely to $\bar{p}$ and $\bar{c}_0$, as defined above. This gives:

Proposition 1 *The average probability estimate in a large economy is almost surely equal to*

$$\bar{p} = \sum_{g=1}^{G} f_g \bar{p}_g, \qquad \textit{where} \tag{14}$$

$$\bar{p}_g = \frac{p + m_g S_\tau(\bar{r})}{1 + m_g}, \;\; g = 1, \ldots, G, \tag{15}$$

and where $\bar{r}$ *as defined in equation (11) is the unique root within the unit interval of the equation*

$$0 = \alpha p + (1-\alpha) S_\tau(\bar{r}) - \bar{r}, \qquad \alpha \stackrel{\text{def}}{=} \sum_{g=1}^{G} \frac{u_g}{1+m_g}. \tag{16}$$

Per capita consumption is almost surely equal to $\bar{c}_0 = \frac{\bar{p}W}{2}$.

The solution to (16) when $\tau > 1$ is

$$\bar{r} = \frac{V + \sqrt{V^2 + 4p\alpha(\tau-1)}}{2(\tau-1)}, \qquad \text{where } V \stackrel{\text{def}}{=} \alpha p(\tau-1) + (1-\alpha)\tau - 1,$$

and the solution when $\tau = 1$ is $\bar{r} = p$. When $\tau = 1$, (15) then immediately implies $\bar{p}_g = p$, for all g, and (14), moreover, implies that $\bar{p} = p$. Agents learn correctly, because in the absence of visibility bias, agents update rationally based on their consumption bin observations. In this case agents are correct in believing that these observations are Ber(p) distributed. Agents of type g therefore use the total $L + M_g$ signals optimally to form beliefs about p.

If $\tau > 1$, agents still believe that the consumption of each type is optimal i.e., that $\bar{p}_g = p$—agents fail to adjust for how visibility bias affects others' behavior. It is easy to verify that when $\tau > 1$,

$$\bar{p} > \frac{p + m_g S_\tau(p)}{1 + m_g}. \tag{17}$$

So there is a positive feedback effect—$\bar{p}$ is higher than the LHS of equation (9), the upward-biased expected probability estimate of an observing agent of type g if all agents consumed based upon the true probability p. In equilibrium an agent has higher consumption owing to visibility bias, which induces higher consumption by other agents. This in turn further encourages high consumption by the original agent.

We call $\bar{p}$ the *equilibrium probability estimate.* It is proportional to per capita consumption, $\bar{c}_0 = \bar{p}\left(\frac{W}{2}\right)$. Since per capita consumption is proportional to the equilibrium probability estimate, throughout the paper we use the two terms interchangeably. The ratio of actual average consumption to rationally optimal consumption is $\frac{\bar{p}\left(\frac{W}{2}\right)}{p\left(\frac{W}{2}\right)} = \bar{p}/p \geq 1$, with strict inequality when there is visibility bias.

Recalling that the levels of observation of different types, m_g, are increasing with g, the next proposition shows that people who engage in greater social observation will overconsume more.

Proposition 2 *In equilibrium $\bar{p}_g$ is increasing in g.*

The result follows directly from (15) and the fact that $S_\tau(\bar{r}) > p$. This implication could potentially be tested with survey data, which has been used to study reported investment behavior in relation to households' sociability or intensity of social interaction, in the form of self-reports of interactions with neighbors or regular church-going (Hong, Kubik, and Stein 2004; Georgarakos and Pasini 2011)). Furthermore, some studies have exploited information about actual social media connections in relation to financial decisions (Heimer (2016), Bailey et al. (2018, 2019)).

3.1 The base model: homogeneous agents

Much insight is gained by studying the case of homogeneous agents, so that there is only one type, $G = 1$, $M_1 = M$, $f_1 = u_1 = 1$, $m_1 = m = M/L$. We call this case the *base model.* In the base model, equations (10)-(12) reduce to the condition

$$\bar{p} = \frac{p + mS_\tau(\bar{p})}{1+m}, \tag{18}$$

with solution

$$\bar{p} = \frac{V + \sqrt{V^2 + 4p(1+m)(\tau-1)}}{2(\tau-1)(1+m)}, \qquad \text{where } V = (p+m)(\tau-1) - 1. \tag{19}$$

The solution has the following properties.

Proposition 3 *In the base model:*

1. *The equilibrium probability estimate, $\bar{p}$, and per capita consumption are increasing in visibility bias, τ, i.e., $\partial\bar{p}/\partial\tau > 0$;*

2. *As visibility bias tends to infinity, $\bar{p}$ approaches $(p+m)/(1+m) < 1$;*

3. *If $\tau > 1$, the equilibrium probability estimate, $\bar{p}$, and per capita consumption are increasing with m, the intensity of observation of others. As $m \to \infty$, $\bar{p} \to 1$, and per capita consumption approaches its maximum possible value, $\bar{c}_0 \to W/2$.*

Proposition 3 generalizes immediately to the case $G > 1$ if $m_g = m$ and $u_g = f_g$ for all g. Under this assumption, agents are still in effect identical.

Part 1 of Proposition 3 says that owing to visibility bias in consumption observations and neglect of sample selection bias (or equivalently, use of the availability heuristic) in assessing bin fullness, people update more strongly toward a belief that others are consuming heavily. In consequence, observers infer too strongly that a wealth disaster is unlikely, which causes them to overconsume. The greater the visibility bias, the larger the effect.[16]

Part 2 indicates that when visibility bias becomes maximally strong, beliefs become maximally overoptimistic, but that agents' private observations have a moderating effect. So equilibrium beliefs do not spiral upward to $\bar{p} = 1$. Agents put some weight on their private observations, so even if 100% of observed bins are full, observers only update their average belief to $(p+m)/(1+m) < 1$. The private signals thus limit the severity of overconsumption.

Part 3 says that owing to visibility bias, greater observation of others, as reflected in m, implies more optimistic beliefs and greater aggregate consumption. As m approaches infinity, the amount of observation of others becomes large relative to the prior precision, and there is drastic overconsumption (agents consume as if they were sure there were no risk of the adverse wealth shock). New biased observations dominate prior information, so that people become certain of a high outcome, even if visibility bias is small ($\tau \approx 1$) and the probability of a high outcome, p, is low.

This is because when agents place heavy weight on socially derived information, the feedback effect becomes very strong. This strong feedback effect also implies that equilibrium consumption is very sensitive to changes in τ for large m.[17] Together, Parts 2 and 3 suggest that the feedback effect inherent in social transmission may be more important in generating severe overconsumption than the direct effect of visibility bias, as long as there is some visibility bias.[18]

Several psychological studies have found that college students overestimate the frequencies of salient behaviors relating to drinking, drug use, and sexual activity.[19] Social psychologists have

[16] Although we do not explicitly derive this, allowing for a few agents with deviant values of τ provides the additional empirical implication that those with higher visibility bias overconsume more than those with lower τ. So empirically, the model implies greater overconsumption on the part of individuals and groups that have lower education and IQ. Psychometric indices such as scores based upon the Cognitive Reflection Task (see the discussion in Frederick (2005)) provide more direct ways of measuring whether an individual is likely to fail to adjust for selection bias (in this case, visibility bias).

[17] When $\tau = 1$, $\bar{p} = p$ regardless of m. When $m \approx 0$, as τ becomes large, updated beliefs are still close to p. In contrast, when m is large, as τ increases from $\tau = 1$, beliefs approach 1.

[18] Empirically, social learning can indeed induce strong feedback effects in consumption behavior. Moretti (2011) provides evidence that social learning about movie quality induces a large 'social multiplier,' wherein observation of others greatly increases the sensitivity of aggregate demand to quality.

[19] For example, studies find that college students overestimate how much other students engage in and approve

argued that when people believe, even mistakenly, that others are engaging in a behavior heavily, they regard the behavior as validated, and engage more in the behavior themselves. Proposition 3 verifies that such an effect does indeed occur in a setting in which agents update their beliefs based upon biased observation of others.

As discussed in the introduction, personal saving rates have plunged in the U.S. and several other OECD countries over the last 30 years, and existing rational theories do not seem to fully explain this phenomenon. Parts 1 and 3 of Proposition 3 provide a possible explanation.

Over the last several decades, improvements in electronic communications by such means as telephone (the drop in cost of long-distance phone service), the rise of cell phones and email in the early 1990s, the rise of internet in the late 1990s, and blogging and social networking (such as Facebook) over the last decade have dramatically reduced the cost of conveying information about personal consumption activities. This is reflected in our model as an increase in both τ and m, as in Parts 1 and 3 of Proposition 3. Greater observation and communication in general about the behavior of others is reflected by higher m in the model. Greater m intensifies the effects of visibility bias by increasing the weight on social observation relative to the prior, and implies a reduction in the savings rate.

Crucially, these technological changes also strongly suggest an increase in bias toward observing consumption over nonconsumption, i.e., visibility bias τ. The activities that are noteworthy to report on very often involve expensive purchases, as with eating out or traveling. Numerous television dramas and reality shows have long had a focus, implicit or explicit, on such consumption activities. The explicit side includes travel and shopping channels. For example, the first national shopping network began in 1985 as the Home Shopping Network. The implicit side includes dramas, not limited to those centered upon the antics of the wealthy ("Who shot JR?"). The shift to reality television also induced greater observation of the consumption activities of others.

In more recent years, social media and review sites have been organized around consumption activities, such as Yelp and TripAdvisor. The universe of YouTube video postings includes travel and other consumption activities. On Facebook, a posting about a consumption event triggers a notification to friends; a non-posting about not engaging in a consumption event does not. Participants in special interest online discussion sites (e.g., focused on high tech or classical music) often post about associated product purchases. Such postings are more interesting, and

of uncommitted or unprotected sexual practices (Lambert, Kahn, and Apple 2003) and heavy alcohol use (Prentice and Miller (1993), Schroeder and Prentice (1998), Perkins and Haines (2005)), and overestimate the use of various other drugs (Perkins et al. (1999)).

therefore more likely to occur, than postings to announce the news that an individual did not buy anything today.

In contrast, in-person unmediated observation of physically proximate friends or acquaintances are likely to often include even nonconsumption activities. So the rise in modern communications results in an increase in visibility bias (i.e., larger τ) and lead to higher overconsumption.[20]

Past social research has also employed other proxies for the intensity of social interaction and observation (m), such as population density (e.g., urban versus rural).[21] This leads to the empirical implication that after appropriate controls, greater population density is associated with lower saving. In the time series, this suggests that the secular increase in U.S. population over time may also have contributed to the decline in the savings rate.

A surprising feature of the base model is that, despite naivete about visibility bias, agents end up with *correct* beliefs about others' average consumption and beliefs. To see why, recall that all agents expect others on average to consume based on the correct value of p. In other words, agents do not recognize that others overconsume. Similarly, each agent thinks that the agent's own consumption is on average based on the correct value p. So each agent believes that the consumption of peers is on average the same as the agent's own. Since all agents are ex ante identical (apart from their unbiased prior signals), they are correct in thinking so. In other words, on average agents correctly assess the average consumption of others.

This may seem counterintuitive, since agents are updating naively about the consumption of others based upon upward-biased samples. However, this is counterbalanced by the fact that on average each agent's prior implies an *underestimate* of others' equilibrium consumption. (Each agent thinks that others are not overconsuming. At the average prior of p, this implies an average belief that others are consuming based upon p. But in equilibrium, other agents on average consume based upon a belief above p.) So people start out with a belief about others that is on average too low, and update too strongly toward a high belief. On average these two effects exactly offset, as must occur by the reasoning in paragraph above.

[20]Increased internet usage—especially through online social networking platforms—is associated with a larger number of 'weak ties' (merely casual acquaintances) in ones' social network. Such weak ties are especially useful for acquiring information and ideas (Donath and Boyd 2004; de Zúñiga and Valenzuela 2011). Also, a social networking platform that relies on advertising for its revenues may have an incentive to disproportionately convey notifications that relate to consumption activities.

[21]People who are geographically closer tend to interact more (Borgatti et al. (2009)), even after the rise of the internet and low-cost telephony (Mok et al. (2010)), and even in online social networks (Scellato et al. (2010)). Sociologists have argued that people in urban areas have more voluntaristic social linkages, as contrasted, e.g., with family ties (White and Guest 2003). These findings suggest that greater population density increases the opportunities for people to interact and observe each other.

3.2 Interpretation as an Observation Network

In a social network environment, the number of other agents observed by a type g agent, M_g, can be thought of as the *in-degree* of an agent of type g in the observation network, so that with L_g independent of g, m_g is proportional to in-degree. On the flip side, $\omega_g \stackrel{\text{def}}{=} u_g/f_g$ can be thought of as being proportional to the agent's *out-degree*, i.e., the number of other agents that observe an agent of type g in the social network. Specifically, ω_g is proportional to how many other agents are followers who potentially receive information from an agent of type g.[22] In other words, linkage between agents is directed.

The following proposition characterizes how the properties of the observation network affect equilibrium consumption:

Proposition 4 *In the model:*

1. *The per capita consumption of any type, $g' = 1, 2, \ldots, G$ and overall per capita consumption are increasing in the in-degree of any type, m_g.*

2. *Consider two economies, A and B that are identical except for the fractions of agents of different types, such that $\{f_g^A\}_{g=1,\ldots G}$ first order stochastically dominates $\{f_g^B\}_{g=1,\ldots G}$. Then per capita consumption is higher in economy A than in B. In other words, higher agent in-degrees promotes aggregate overconsumption.*

3. *Consider two economies, A and B, that are identical except for the out-degrees of different agent types, such that $\alpha^A < \alpha^B$, where $\alpha^A = \sum_{g=1}^{G} \frac{f_g}{1+m_g}\omega_g^A$, $\alpha^B = \sum_{g=1}^{G} \frac{f_g}{1+m_g}\omega_g^B$. Then per capita consumption is higher in economy A than in B, i.e., greater association of out-degree with in-degree across types promotes aggregate overconsumption.*[23]

Part 1 shows that higher in-degree on the part of agents of a given type always increases per capita consumption in society at large. The effect on consumption also applies to the effect upon any type, i.e., for any g, g', an increase of in-degree among type g agents increases per capita consumption of type g'. Intuitively, owing to neglect of visibility bias, high in-degree biases agents' beliefs more heavily in favor of high consumption. Their high consumption has a positive

[22]Each agent of type g' observes $M_{g'}$ other agents, and the fraction $f_{g'}$ are of type g'. So in total $N\sum_{g'} M_{g'}f_{g'} \stackrel{\text{def}}{=} N\kappa$ observations are made by agents of all types. The parameter κ represents the number of observations per capita. Fraction u_g of these are observations of group g's consumption bins. So in total there are $u_g N\kappa$ such observations from the bins of the $f_g N$ agents of type g. The average number of observations made (per capita in the population) of an agent in group g is therefore $u_g N\kappa/(f_g N) = \omega_g\kappa$. In other words, ω_g is proportional to the out-degree of group g agents.

[23]To see that α is a measure of association, observe that $\alpha = \sum_{g=1}^{G} \frac{u_g}{1+m_g} = \widehat{cov}(v, (1+m)^{-1}) + \overline{v}\,\overline{(1+m)^{-1}} = \widehat{cov}(v, (1+m)^{-1}) + \overline{(1+m)^{-1}}$. So α is a measure of negative covariation of v and m across types.

feedback effect on the consumption of others. A similar intuition applies to Part 2, which shows that when the population is more heavily tilted toward types with high in-degree, per capita consumption is higher.

Part 3 provides a precise measure of how out-degree affects consumption. It implies that overconsumption is especially severe in social networks where agents with high in-degrees (m_g) are also those with high out-degrees (u_g). Intuitively, agents with high in-degree are more biased toward high consumption. When such agents also have high out-degree, this high consumption is disproportionately influential in encouraging other agents to consume heavily. As discussed earlier, it is likely that high in-degree is positively associated with high out-degree, since in reality people differ in overall sociability, which promotes both observation of others and being observed by others. Krapivsky, Rodgers, and Redner (2001) provide a model of network formation with this property. This effect further amplifies the tendency toward overconsumption identified in the base model.

An implication of Proposition 4.1 is that if a subset of the population, through trends such as the rise of low-cost travel, electronic communication, and social media, become more heavily connected, overconsumption will increase, in aggregate, even among groups that do not increase their social connections.

3.3 Policy Interventions

Overconsumption in our model derives from underestimation of vulnerability to adverse wealth shocks (overestimation of p). This suggests that a relatively simple policy intervention—saliently publicizing valid information about the risk of wealth shocks—can help alleviate overconsumption. For example, publicity about the frequency of layoffs or of expensive illness could be beneficial. Interpreted more broadly, low p could be the risk of living a long time, resulting in higher-than-expected post-retirement consumption needs. So salient publicizing of life expectancy information could help.

However, in practice, such disclosure may be non-salient and hard for people to interpret. Research on heuristics and biases consistently finds that people tend to put little weight on base rate frequency or probability information (Kahneman and Tversky 1973; Borgida and Nisbett 1977). So a numerical report about frequency of layoffs is likely to have little effect. Furthermore, people may have trouble interpreting mortality or life-expectancy tables, which require significant cognitive processing to translate into an optimal plan for how much to save. We therefore consider other possible types of disclosure.

A possible policy intervention suggested by the visibility bias approach is to publicize something simple which translates fairly directly into a consumption/savings recommendation: the average consumption or saving rate of peers. Under plausible variations of the base model assumptions, we shall see that people can end up with biased perceptions about what others believe and how much they consume. If so, saliently publicizing accurate information about peers can help alleviate overconsumption. Specifically, if people overestimate how optimistic their peers are, and how much others consume, then accurate information about others would correct that mistake, reducing overconsumption.[24]

Empirically, in tests covering a very wide range of activities, people often misperceive the beliefs or behaviors of peers, and the intervention of providing accurate information about peers tends to cause behavior to conform more closely to the disseminated peer norm.[25] So "social norms marketing," or dissemination of information about the behavior or beliefs of one's peers, can be an effective policy tool for correcting inaccurate beliefs about peers, and to makes peer actions more salient.

We now extend the model to allow for unbiased public information disclosure about the consumptions of other agents. Specifically, the disclosure provides each agent n with Q visibility-*unbiased* i.i.d. signals, z_k^n , $k = 1, ...Q$, about the equilibrium consumption of other agents. We denote by v_g the fraction of these Q signals that come from type g agents (a fraction that does not depend on the agent n).[26]

When the fraction of visibility-unbiased signals that are about type g is in proportion to population share, $v_g = f_g$ for all g, the new signals may be viewed as being about per capita consumption, $\bar{p} = \frac{1}{Q}E[\sum \tilde{z}_k^n]$. We also allow for other signal weights, since the policy maker may

[24] This raises the question of what the relevant set of peers is, and how people recognize peers. In the context of our model, in which agents are learning from others about the risk of wealth disaster, peers are agents who have identically distributed wealth shocks. People probably recognize peers via geographical proximity (neighbors), social and professional relationships (coworkers and people at a similar professional or socioeconomic level that they transact with) and extended family.

[25] See e.g., Schroeder and Prentice (1998). Frey and Meier (2004), Cialdini et al. (2006), Salganik, Dodds, and Watts (2006), Goldstein, Cialdini, and Griskevicius (2008), Cai, Chen, and Fang (2009), Gerber and Rogers (2009), Chen et al. (2010) and Bursztyn, González, and Yanagizawa-Drott (2020). However, there are exceptions in which people adjust their behavior away from the disclosed actions of others. In an experiment on retirement savings behavior in a large manufacturing firm, Beshears et al. (2015) document that information about the high savings rates of other employees can sometimes lead low-saving individuals to shift away from the disclosed savings rates, which Beshears et al. suggest may derive from a discouragement effect. This result holds only for the subpopulation of employees with low relative incomes who had never participated in the firm's 401(k) plan. Such employees may regard the higher-income employees who were plan participants as not truly peers (e.g., in our model, having non-identically distributed wealth shocks). So our theory does not make a prediction about the outcome of this experiment.

[26] We can alternatively think of this signal structure as representing a common public disclosure signal, about the weighted average consumption, $\sum_g v_g \bar{p}_g$, which agents observe with agent-specific noise. Then the higher is Q, the lower is the noise.

be able to provide a disclosure focused upon the consumption of specific groups (where everyone accurately understands this focus).

Since agents believe that the average consumption of each type is optimal, i.e., that $\bar{p}_g = p$ for all g, they treat each additional signal provided by the policymaker as being exactly as informative as the other signals agents receive. By an argument similar to the basic analysis, we obtain the following updating rule for type g agents:

$$\hat{p}^n = \frac{\bar{x}^n + m_g S_\tau(\bar{y}^n) + q_g \bar{z}^n}{1 + m_g + q_g}, \tag{20}$$

where $\bar{z}^n \stackrel{\text{def}}{=} \frac{1}{Q}\sum_{k=1}^{Q} \tilde{z}_k^n$, and $q_g = Q/L_g$. Proposition 1 can now be generalized to allow for disclosures about agents' consumptions.

Proposition 5 *The average probability estimate in a large economy in which agents observe additional disclosure signals of other agents' equilibrium consumption is almost surely equal to*

$$\bar{p} = \sum_{g=1}^{G} f_g \bar{p}_g, \qquad \textit{where} \tag{21}$$

$$\bar{p}_g = \frac{p + m_g S_\tau(\bar{r}) + q_g \sum_g v_g \bar{p}_g}{1 + m_g + q_g}, \quad g = 1, \ldots, G, \tag{22}$$

and where $\bar{r}$ is the root within the unit interval of the equation

$$0 = \beta p + (1 - \beta) S_\tau(\bar{r}) - \bar{r}, \tag{23}$$

where

$$\beta \stackrel{\text{def}}{=} \sum_{g=1}^{G} \frac{u_g}{1 + m_g + q_g} + \frac{1}{1 - \sum_{g=1}^{G} \frac{v_g q_g}{1 + m_g + q_g}} \left(\sum_{g=1}^{G} \frac{u_g q_g}{1 + m_g + q_g} \right) \left(\sum_{g=1}^{G} \frac{v_g}{1 + m_g + q_g} \right). \tag{24}$$

The per capita consumption is almost surely equal to $\bar{c}_0 = \frac{\bar{p}W}{2}$.

Equation (22) provides a linear system of equations for $\bar{p}_g$, which can easily be solved once $\bar{r}$ is known. The general model we have studied so far corresponds to the case when $Q = 0$, in which case $\beta = \alpha$, where α is defined in (16).

For tractability, we study the special case of two groups, $G = 2$. The in-degrees and number of private observations can differ for groups 1 and 2, with $m_1 = M_1/L_1$ and $m_2 = M_2/L_2$, so that some agents are more prone than others to be influenced by biased social observations. The out-degrees $\omega_1 = u_1/f_1$ and $\omega_2 = u_2/f_2$ can also differ, so that some agents are more influential than others. Specifically, agents of a type with a very high L_g will have a very precise estimate of p from private signals, and will only change this estimate marginally when incorporating social consumption bin (and disclosure) observations. We think of such agents as being "smart."

Proposition 6 *In the model with public disclosure and with $G = 2$ types:*

1. *A sufficient condition for per capita consumption to be decreasing in the amount of disclosure, $\frac{\partial \bar{p}}{\partial Q} < 0$, is that ω_1 be sufficiently high,*

$$v_1 > \bar{v} \stackrel{\text{def}}{=} \frac{1}{\frac{L_1(1+m_1)}{L_2(1+m_2)}\left(\frac{1}{f_1 \max(1,\omega_1)} - 1\right) + 1}.$$

2. *A sufficient condition for per capita consumption to be increasing in the amount of disclosure, $\frac{\partial \bar{p}}{\partial Q} > 0$, is that ω_1 be sufficiently low,*

$$v_1 < \underline{v} \stackrel{\text{def}}{=} \frac{1}{\frac{L_1(1+m_1)}{L_2(1+m_2)}\left(\frac{1}{f_1 \min(1,\omega_1)} - 1\right) + 1}.$$

3. *When $L_1 = L_2$, per capita consumption is decreasing in the consumption disclosure signal's weight on agents with low equilibrium consumption, v_1, i.e., $\frac{\partial \bar{p}}{\partial v_1} < 0$.*

Proposition 6 has several immediate implications.

Corollary 1 *In the extended model with public disclosure and $G = 2$ types:*

1. *If disclosure is solely about the consumption of type 1 agents, $v_1 = 1$, then per capita consumption is decreasing in the amount of public disclosure, $\frac{\partial \bar{p}}{\partial Q} < 0$.*

2. *If disclosure is solely about the consumption of type 2 agents, $v_2 = 1$, per capita consumption is increasing in the amount of public disclosure, $\frac{\partial \bar{p}}{\partial Q} > 0$.*

3. *If disclosure is about per capita consumption, i.e., $v_1 = f_1$, then:*

   (a) *When the number of private observations is the same for both groups, $L_1 = L_2$, as is out-degree, $\omega_1 = \omega_2$, then per capita consumption is increasing in the amount of disclosure, $\frac{\partial \bar{p}}{\partial Q} > 0$.*

   (b) *When the number of private observations for type 1 agents is sufficiently high,*

$$\frac{L_1}{L_2} > \frac{1+m_2}{1+m_1} \max\left(\frac{1/f_1 - 1}{1/u_1 - 1} - 1, 1\right),$$

per capita consumption is decreasing in the amount of disclosure, $\frac{\partial \bar{p}}{\partial Q} < 0$.

Several conclusions can be drawn from Proposition 6 and Corollary 1. First, Corollary 1 Parts 1 and 2, and Proposition 6 Part 3 jointly suggest that if possible, the policy maker should focus on disclosing the consumption of low-consumption agents. This is a straightforward consequence of the neglect of selection bias in our approach. Since agents believe that all types have the same consumption, they revise down their estimates when being informed about agents with low consumption.

Proposition 6 Parts 1 and 2 provide sufficient conditions for consumption to be decreasing or increasing in the amount of disclosure. Both conditions are feasible, $0 < \underline{v} < \bar{v} < 1$, so disclosure can either increase or decrease consumption, depending upon how heavily the disclosure signal is weighted toward type 1 versus type 2 agents. In comparison, for the base model, with $G = 1$, $v_1 = \underline{v} = \bar{v} = 1$, so disclosure never affects consumption.

Focusing on the case when all agents make the same number of private observations ($L_1 = L_2$), Parts 1 and 2 of Proposition 6 show how agents' in-degrees (measured by m_g), out-degrees (measured by ω_g) and the fraction of each type (measured by f_g) jointly determine how much weight needs to be put on the low-consumption agents of type 1 in the disclosure signal, v_1, for public disclosure to reduce overconsumption.

If all agents have the same out-degree, so that $\omega_1 = \omega_2 = 1$, we are in the benchmark case in which the bound is tight, $\underline{v} = \bar{v}$. In this case, the higher the fraction of low-consuming type 1 agents, the more weight needs to be put on low-consuming agents in the disclosure signal for consumption to decrease (i.e., the higher v_1 needs to be), because these agents' consumptions are already well-reflected in the public observations. A low ratio of in-degrees between the two types, $\frac{1+m_1}{1+m_2}$ lowers the threshold for reducing overconsumption, by making the lower consumption of type 1 agents more surprising for the visibility biased population, when disclosed.

When the out-degree of type 2 agents is higher than that of type 1 agents, $\omega_2 > 1 > \omega_1$, the weight on type 1 agents needed in the disclosure signal to ensure that overconsumption is reduced increases compared with the benchmark case, i.e., $\bar{v}$ increases. Intuitively, it becomes harder to offset the disproportionate influence of the high consumers through public disclosure, so even more weight needs to be put on the low-consuming agents. The effect is the opposite when the out-degree of type 1 agents is higher than of type 2 agents, $\omega_1 > 1 > \omega_2$, for which $\underline{v}$ is lower than in the benchmark case.

Corollary 1 Part 3(a) focuses on the case when disclosure is about per capita consumption, and where there is neither heterogeneity in agent "smartness" nor in out-degree. In this case, disclosure unambiguously increases per capita consumption. The reason for this result is that agents with lower-than-average consumption will tend to increase their consumption when seeing the disclosure, whereas agents with higher-than-average consumption will tend to decrease theirs. When $L_1 = L_2$, the lower-than-average consumption agents are more influenced by the disclosure signal than the higher-than-average consumption agents, since the reason why they consume less than average is that they have relatively few social consumption bin observations, i.e., they have low perceived belief precision. Overall, per capita consumption increases.

In contrast, Corollary 1 Part 3(b) indicates that when agents chose low consumption because they are very smart, i.e., because $L_1 \gg L_2$, disclosure decreases per capita consumption. The reasoning is the opposite in this case. The smart low-consumption agents, being confident about the optimal consumption, barely revise their probability estimates upward when observing the disclosure signal, whereas the "dumb" high-consumption agents substantially revise their estimates substantially downward. So overall per capita consumption decreases.

So a key determinant of whether disclosure of per capita consumption is helpful in reducing consumption, is whether agents who consume less than others have higher or lower subjective confidence about their beliefs than agents who consume more than others. If a set of agents consumes less than another set because the low consumers have greater genuine knowledge (observation of many unbiased signals), the low consumers will have higher confidence in their beliefs. In contrast, if a set of agents consumes less than another set because the low-consumers make few social observations, the low consumers will have lower confidence in their beliefs.

So empirically, the predicted effect of disclosure of aggregate consumption depends on the dispersion in the population of genuine information versus the dispersion in social observation. If agents differ mostly in how genuinely informed they are (number of nonsocial signals), then the first effect is the relevant one, and we expect disclosure to reduce consumption. If instead agents differ mostly in how heavily they observe others (number of social signals), then the second effect is the relevant one, and we expect disclosure to increase consumption.

A possible proxy for agents being privately well-informed is level of education or socioeconomic status. A possible proxy for number of social signals observed is social connectedness, which has been measured in empirical studies using survey proxies, population density, or social media datasets. Our model implies that such data, along with consumption data, can be exploited to predict whether disclosure with increase or decrease per capita consumption.

The effect of disclosure has been tested by D'Acunto, Rossi, and Weber (2020) in a field experiment using a smartphone app. They disclose the average of (income-normalized) spending of other subjects to *overspenders* (those with above-average spending) and *underspenders* (those with below-average spending). This disclosure causes overspenders to decrease spending on average by 3%, whereas underspenders increase their spending by 1%. The result that low-spenders adjust up less than high-spenders adjust down is in line with the smart agent effects of disclosure in our model. Our model implies that when disclosure has this asymmetric effect, underspenders are more knowledgeable or sophisticated than overspenders. This is potentially testable by examining the educational or financial literacy status of these two groups.

Later, in our extension to a setting with overlapping generations, we show that a different kind of disclosure of consumption of a subset of the population—older agents—can also help reduce overconsumption.

The friendship paradox in social networks is the fact that the average number of friends (connections) of agents' friends is higher than the average number of friends agents have. This occurs because well-connected agents are friends of more agents. In some settings, the friendship paradox can cause well-connected agents to be disproportionately influential. As discussed in the introduction, Jackson (2019) examines a setting in which, owing to naivete about the friendship paradox in combination with strategic complementarities (which are not present in our model), people misestimate the average actions of others. As a result, beliefs and aggregate actions can sometimes be corrected by public disclosure. Even in our network setting, where we allow for effects that relate to the friendship paradox, our model differs in deriving implications for general consumption/savings levels without the requirement of strategic complementarities.

3.3.1 Visibility-Biased Disclosures

We now examine disclosure that is subject to an opposite visibility bias: it increases the salience of *non*consumption/saving relative to consumption. We first show that accurate public disclosures that are visibility-biased toward nonconsumption encourage saving. This effect does not require the agent heterogeneity considered in the preceding subsection, so we return to the base model in which all agents are ex ante identical. We then discuss how disclosure in practice can highlight saving behavior.

Consider an accurate public information disclosure that calls attention to saving, i.e., calls attention more to empty consumption bins than to full bins. Just as before, the signal is accurate—it never misrepresents whether a disclosed bin is empty or full.

For example, people could be given stickers to post on their cars or in personal spaces, saying "Proud Retirement Saver." The policymaker could publicize that these signs or stickers are given to anyone who is saving more than some prespecified absolute amount or fraction of income. The visibility bias derives from the greater salience of the presence of the sign or sticker than the absence of one.[27] This induces visibility bias toward observation of heavy contributions.

In the model, if there is visibility bias in the disclosure toward the occurrence of saving behavior, and observers neglect this visibility bias, then investors update strongly toward a belief that others are saving heavily. For presentational simplicity, we assume that the visibility bias toward observing empty bins in the disclosed signal has the same value, τ, as the visibility bias toward observing full bins in the direct observation of others. In other words, for the disclosed signal about saving, the same visibility bias function S_τ is applied, but it is applied to the amount saved rather than the amount consumed. It follows that the equilibrium average belief satisfies

$$\bar{p} = \frac{p + mS_\tau(\bar{p}) + q(1 - S_\tau(1 - \bar{p}))}{1 + m + q}. \tag{25}$$

Here Q savings signals are disclosed and $q \stackrel{\text{def}}{=} Q/L$. The $(1 - S_\tau(1 - \bar{p}))$-term in this expression reflects an agent's neglect of visibility bias about saving. The agent believes that the agent is observing a signal about $1 - \bar{p}$ (proportional to saving, i.e., the fraction of empty bins), when the agent is actually observing a signal about the visibility-biased quantity $S_\tau(1 - \bar{p})$. This leads to the observer to infer that the average belief (proportional to consumption) of those observed via the disclosure is $1 - S_\tau(1 - \bar{p})$, rather than $\bar{p} = 1 - (1 - \bar{p})$.

Equation (25) has a unique closed form solution, and it is not hard to derive that:

Proposition 7 *Per capita consumption when there is visibility bias toward a disclosure of higher saving, as given in (25), is lower than when there is no disclosure, as given in (18).*

Disclosure of a signal that is visibility-biased toward saving fights visibility bias with visibility bias. Visibility biased disclosure encourages saving, thereby partly offsetting the basic tendency toward overconsumption of Proposition 3.

For several reasons, interventions by policymakers to make saving behavior more salient, as reflected in the q term in the numerator of equation (25), is unlikely to fully offset the spontaneous visibility bias toward observing the consumption activities of others, as reflected in the m term. People are heavily and frequently exposed to the consumption of others as people interact with

[27] This procedure provides little incentive to post the stickers, and some people may want to avoid the appearance of bragging. A more nuanced scheme might be localized by neighborhood. Locations could be established for neighbors to post their stickers anonymously in addition to possible posting on cars, homes or at the workplace. The policymaker publicizes a count of the number of stickers posted on the public sign locations, or a count of total self-reports about meeting the saving criterion. Savers as a group are given monetary bonus when the neighborhood count is higher.

others in person or electronically. In contrast, policy campaigns are likely to be episodic and observed by their targets only occasionally. Nevertheless, if sufficiently salient, disclosure may have a beneficial effect.

4 Extensions

To address the generality of our conclusions and to examine additional issues, we now consider variations of the base model. For tractability we assume that $G = 1$, though we allow for agent heterogeneity in other ways; and, as needed, make some stronger assumptions.

First, in reality people observe others at different life-cycle stages. Does observation by the young of low consumption by seniors alert the young to the dangers of consuming too heavily, undermining the conclusion that the young overconsume? Also, does a bias toward observing the young versus the old have empirical implications for how heavily people consume? We address these topics in Subsection 4.1.

Second, in reality people do not perfectly know each other's wealths. This adds noise to the learning problem, because observed consumption is influenced by the wealths of observation targets, not just their information signals about the probability of not experiencing a wealth disaster. This suggests that wealth dispersion may affect overconsumption, a topic we address in Subsection 4.2.

Third, we have so far assumed an exogenous riskfree interest rate. We allow for increasing supply of credit as a function of the interest rate in Subsection 4.3. We show that overconsumption is obtained in this setting too, and that interest rates are higher when visibility bias is present than when it is not.

As a matter of robustness, we also show in the appendix that results similar to those in the base model arise under two technical variations. We verify that similar results apply with other possible utility functions in Appendix A.

We then depart from the assumption that the maximal fraction of full consumption bins is 100% in Appendix A. Specifically, our base model made the assumption that when the agent is maximally optimistic, and therefore consumes $W/2$ at date 1, that this is achieved when all the bins are full. Our extension allows for consumption of $W/2$ to leave some bins empty. This variation is also useful for the analysis in Sections 4.2 and 4.3.

For the variations we study in this section, we make the additional assumption that the numbers of private signals and consumption observations, L and M, are large, so that agents' priors are very close to p and the fraction of observed full bins, $\bar{y}^n$, is very close to $E[\tilde{y}^n]$. The fraction $m = M/L$ is still an arbitrary positive number. We can think of this as studying the limit of a sequence of economies as $L \to \infty$, with $M = mL$ in each economy.

4.1 Age Differences: An Overlapping Generations Setting

In the base model, all agents observe each other and make their savings decision at the same time, when young. What if young people observe the consumption of seniors who are consuming from their savings? If the young overconsume, then as in the base model, observations of the young promote an inference of low disaster risk, which encourages the young to consume heavily. However, if the young overconsume, the old, on average, underconsume. If young observers see *low* consumption of the old (owing to visibility bias, this may not be the case), the observer may infer that disasters were realized heavily, reducing consumption. This implies an inference by young observers of *high* disaster risk. This raises the question of whether the young will actually overconsume.

An alternative perspective also raises possible doubt about the prediction that the young overconsume. Suppose that owing to visibility bias, young observers think they see *high* consumption by old agents. Then young agents may infer that old agents had saved a lot when they were young, which would occur if they had viewed the risk of a wealth disaster as high. This inference discourages young observers from consuming heavily.[28]

Fortunately, explicit modeling brings clarity. When modelled in a straightforward way, just as in the base model, unambiguously, the young overconsume. To allow for observation of the old, we extend the base model to include an overlapping generations (OLG) structure in which there are both young and old agents at any given point in time.

We now assume that the ϵ-shock is independent across agents (though still identical in distribution), which rules out fluctuations in aggregate consumption across time. Moreover, we study a stationary equilibrium in which the average estimated probability of no wealth disaster, $\bar{p}$, is constant over time.

We also assume that fraction $\lambda \in [0, 1]$ of the bin observations are of the young, and the remaining fraction $1 - \lambda$ of observations are of the old. So young agents observe a random sample of consumption from each cohort, i.e., λM observations are from the young generation, and $(1 - \lambda)M$ from the old.

[28] Alternatively, a young observer might conclude that in the current period old agents have generally had favorable wealth realizations (not disasters). That suggests an inference of low disaster risk, which encourages young agents to consume heavily. The overall outcome is not immediately obvious.

The case $\lambda = 1$ corresponds to the base model. The young might, for example, disproportionately observe each other rather than the old, owing to homophily (the tendency for people to interact with others who are similar), leading to higher λ. On the other hand, the old may act as role models for the young, leading to lower λ.

In addition to possible bias toward observing young or old, λ reflects the fractions of the population that are in these two groups. If, for example, there is no bias in observation of young versus old, we can think of λ as capturing population growth, with λ high in rapidly growing populations. Alternatively, it can inversely capture longevity, since long lifespan increases the fraction of the population that is old.

Introducing observation of the old requires a slight extension of the base model to address a technical issue. A senior who is unlucky and hit with disaster potentially has a negative level of consumption, but it does not make sense to talk about a negative fraction of full consumption bins. Likewise, a senior who is lucky and not hit with a disaster potentially consumes more than $W/2$, the expenditure that corresponds to all bins being full. Since the reasoning of the model is based on average levels of consumption within the population or subpopulations, we address these issues by assuming a date 1 transfer of consumption from the lucky old to the unlucky old that brings the consumption level of the unlucky up to zero, and ensures that the consumption levels of the lucky old are never greater than $W/2$.[29]

We assume that visibility bias, as previously specified in (7), is the same for observations of either the young or the old. As before, observers think there is no visibility bias, and believe that the average consumption of the young is optimal, $pW/2$, where as before, p is the true probability of non-disaster. It follows that observers believe that the average consumption of the old is also $pW/2$, corresponding to the fraction p of full consumption bins by the old.[30] The young therefore view consumption bin observations of the old as having identical information content as observations of the young.

A consequence of this is that it does not matter for the analysis whether observers can see whether any given consumption bin observation is drawn from the old or from the young. Even if an observer can see the identity of the agent corresponding to an observed bin, the observer

[29]Specifically, at date 1, unlucky seniors work for lucky seniors (e.g., shopping for them or mowing their lawns), the payment thereby increasing the consumption level of the unlucky to zero, and reducing the average consumption of the lucky seniors accordingly. We further assume that the utility gains from this exchange are zero, so that the disutility of work of the unlucky offsets their consumption benefit, and the utility benefit to the lucky of hiring the unlucky is also zero. Since these transactions leave everyone indifferent ex post, the ex ante optimization problem at date 0 is unchanged.

[30]The average consumption of old who consumed $pW/2$ when young is, by the law of large numbers, $p\,(W - pW/2) + (1-p)(W - pW/2 - W) = pW/2$.

ignores the identity information.

Given a cohort's estimated $\bar{p}$ when young and their associated consumption of $\bar{p}\left(\frac{W}{2}\right)$, by the law of large numbers their average consumption in the next period, when old, is $(2p-\bar{p})\left(\frac{W}{2}\right)$, where p is the true probability. When $\bar{p} > p$, there is underconsumption by the old generation compared with the social optimum, $pW/2$, since $2p - \bar{p} < p$.

Owing to visibility bias, observation of the bins of the young and the old are biased toward full bins, as reflected in the S_τ function. So by reasoning very similar to that leading to (6), the equilibrium condition is

$$\bar{p} = \frac{p + mS_\tau[\lambda\bar{p} + (1-\lambda)(2p-\bar{p})]}{1+m}. \tag{26}$$

In the benchmark case of no visibility bias ($\tau = 1$), in equilibrium $\bar{p} = p$ and both cohorts consume on average $pW/2$. In this case, observations of young and old consumption are consequently equally informative, and agent beliefs and behavior are rational.

Proposition 8 *In the OLG extension of the base model, there is a stationary equilibrium satisfying the following properties:*

1. *The equilibrium probability estimate of the young generation satisfies $\bar{p} > p$, so the young generation overconsumes.*

2. *The equilibrium probability estimate, $\bar{p}$ is increasing in the fraction of young agents, λ.*

3. *When $\lambda = 0$,*

$$\begin{aligned} \bar{p} &= \frac{1 + m(\tau+1) + p(3+2m)(\tau-1) - \sqrt{V}}{2(1+m)(\tau-1)}, \qquad \textit{where} & (27)\\ V &= (1 + m(\tau+1) + p(3+2m)(\tau-1))^2 \\ &- 4p(1+m)(\tau-1)(1+2p(\tau-1)+2m\tau). & (28) \end{aligned}$$

An implication of Proposition 8 Part 2 is that overconsumption is greater in economies with rapid population growth, low longevity, or in which observation is more heavily tilted toward the young. It is of course crucial to have appropriate controls (e.g., for investment opportunities) in cross-economy tests.

Proposition 8 shows that the young unambiguously overconsume in equilibrium, even when the young predominantly observe the old, or even entirely do so ($\lambda = 0$). This may seem surprising given the mixed intuitions at the start of this subsection. The first intuition started from the fact that if the young on average overconsume, then the old on average underconsume. So a tilt of observation toward the old tilts the inference toward low p (high disaster risk), favoring underconsumption. This suggests that heavy overconsumption by the young may not be self-confirming.

Nevertheless, this effect cannot reverse equilibrium overconsumption by the young. If, out of equilibrium, the amount of overconsumption by the young were arbitrarily small, then the average consumption of the old would be almost as high as average consumption of the young. Visibility bias in observation of others—even of the old—would then result in a high inference about the consumption of others, which non-negligibly favors overconsumption.

The alternative argument given at the start of this subsection for why there might be underconsumption was that owing to visibility bias, observers think they see seniors consuming heavily. Such apparent high consumption by seniors might be taken to mean that seniors, when young, had adverse information about the wealth shock. Why doesn't this lead young observers to conclude that they need to save heavily rather than consume heavily?

The flaw in this argument is that observers believe that the old, when young, were on average consuming optimally, i.e., consuming $pW/2$, which implies that the average consumption of the older generation when old is the same, $pW/2$. In other words, the young think that the probability that any observed bin is full (regardless of whether it is drawn from a young or old agent) is p. So a full bin is always indicative of a *high* probability of non-disaster. So visibility bias in observations of others, old as well as young, encourages *high* consumption.

This insight makes clear why in equilibrium, overconsumption is greater when observation is more heavily tilted toward the young (λ high). Relative to observation of the young, observation of the old acts as a partial reality-check on belief bias. Observers mistakenly think that on average consumption is equally divided between an agent's youth and old age, but owing to overconsumption, in equilibrium, *actual* average consumption is lower for old agents. So observations of the consumption bins of the young are more often full than observations of the bins of the old. So higher λ (sampling from the young) leads to more favorable inferences about p, and therefore greater overconsumption.

This reasoning provides insight into a possible objection to the basic implication of our approach that visibility bias induces overconsumption. The purchase of a house could be an indicator that an individual had saved heavily to accumulate enough funds for a substantial down-payment. It might be argued that others will infer from this that it is important to save heavily. However, as the preceding paragraph makes clear, when young agents observe high consumption (including housing consumption) by seniors, there is a favorable inference about disaster probability, resulting in overconsumption.[31]

These findings have an implication for optimal disclosure policy—that salient public disclosure of the consumption of the old can help reduce overconsumption. This intervention consists of disclosing the average consumption of a *subset* of the population. The effect of such disclosure is effectively to push the model in the direction of low λ (in which there is more observation of consumption of the old). As we have shown, lower λ decreases aggregate consumption. It is interesting that even in scenarios where disclosing aggregate consumption of the entire population does not help, disclosing the consumption of the right subset does help. However, such disclosure does not fully remedy the problem. At best it only reduces overconsumption to that of the $\lambda = 0$ case.

4.2 Uncertainty about Wealth

So far, we have assumed that all agents have the same non-disaster wealth level W. We now generalize to allow for ex ante wealth dispersion in the population (even apart from ex post disaster realizations), and ignorance of the wealths of others. Intuitively, the inference an observer draws about others' signals based upon observing others' consumption is diluted by ignorance of the wealth of the observation target. High apparent consumption of an observation target could come either from the target possessing a favorable signal (indicating low risk of disaster), or from the target having higher ex ante wealth. Observers will therefore not, on average, revise their probability estimate $\hat{p}$ upward as aggressively as they do when there is no information asymmetry about wealth. Wealth uncertainty (and wealth dispersion associated with such uncertainty) therefore reduces equilibrium overconsumption. This contrasts sharply with the Veblen wealth-signaling approach discussed in the introduction, in which it is uncertainty about wealth that is the source of overconsumption signaling.

To allow for wealth uncertainty and dispersion, we now assume that a fraction λ (not the λ of the preceding section) of the population has non-disaster wealth level $(1 + \Delta)W$ (the wealthy group), where $\Delta > 0$, a fraction λ has non-disaster wealth $(1 - \Delta)W$ (the poor group), and the remaining fraction $1 - 2\lambda$ has non-disaster wealth W (the medium group). Henceforth we refer to "non-disaster wealth" more briefly as "wealth." The average wealth is then still W, but the

[31]Buying a house is an example of the general high visibility of engaging in consumption activities. The purchase of a house is usually a shift to a higher flow of current consumption of housing services financed by a major increase in indebtedness (mortgage down payments are usually much smaller than the size of the loan). Indeed, real estate equity is often accessed to finance non-housing consumption expenditures as well (Chen, Michaux, and Roussanov 2019). Lusardi, Mitchell, and Oggero (2018) report that in recent years, older Americans close to retirement hold more debt than earlier generations, primarily owing to the purchase of more expensive homes with smaller down payments.

higher Δ is, the higher the wealth uncertainty and the wealth dispersion in the economy.[32] We continue to assume that there is a probability $1-p$ that any given agent (rich, poor, or medium) experiences a disaster (negative value of ϵ) that entirely wipes out an agent's wealth at date 1.

So far, we have assumed that when the agent consumes half of potential wealth at date 0, $c_0 = W/2$, that all consumption bins are full. To explore the effects of wealth dispersion,, we now need to allow for the possibility that someone with maximally optimistic beliefs, $\hat{p} = 1$, still has some empty consumption bins. We now instead assume that when $\hat{p} = 1$,, only a fraction $0 < h \leq 1$ of the bins are full. To prevent the wealthy group from consuming so heavily that more than 100% of the consumption bins are full, we assume that $(1+\Delta)h < 1$. In addition, we impose the technical condition that

$$\Delta \leq \frac{1}{1+\frac{2}{\tau-1}}. \tag{29}$$

For large τ, this implies $\Delta < 1$. Since poor agents have wealth $(1-\Delta)W$, for large τ this imposes the mild restriction that even poor agents do not have wealth very close to zero.

Agents know the economy's wealth distribution (λ and Δ), and for simplicity we assume that each agent's consumption bin observations come from one or more agents of the same wealth type (where the agent does not know which type). Based on these observations, an agent forms a posterior belief about p, taking into account that an observation of high consumption could reflect high wealth, not just an optimistic belief, on the part of the target of observation.

The following proposition confirms the intuition that wealth uncertainty reduces overconsumption:

Proposition 9 *Under the above assumptions, the equilibrium probability estimate $\bar{p}$ and overconsumption are decreasing in wealth uncertainty, $\partial \bar{p}/\partial \Delta < 0$.*

Proposition 9 predicts that savings rates increase with wealth uncertainty. This is the opposite of what is expected based upon Veblen wealth-signaling considerations.

Specifically, in the Veblen approach to overconsumption, people consume more in order to signal the level of wealth to others (Bagwell and Bernheim 1996; Corneo and Jeanne 1997), so there is no motive to overconsume when wealth uncertainty is zero. A comparison of the cases of zero versus positive wealth uncertainty indicates an average tendency for greater wealth uncertainty to induce greater overconsumption, though not necessarily monotonically.[33] This

[32] We do not focus upon learning about mean wealth in the population. Extending our setting to allow for learning about mean population wealth (in the absence of wealth uncertainty) would not affect agents' decision problems, since each agent is concerned about the probability of disaster, not about how wealthy others are per se.

[33] Consistent with this idea, Charles, Hurst, and Roussanov (2009) find empirically that greater dispersion in

intuition is reflected, for example, in a comparative statics of Charles, Hurst, and Roussanov (2009) in which a parameter shift that increases wealth uncertainty by reducing the lower support of wealth results in greater consumption signaling.[34] Furthermore, and also in contrast with our approach, in at least one version of the keeping-up-with-the-Joneses approach, income dispersion encourages the non-wealthy to consume more in emulation of the wealthy (Bertrand and Morse 2016).[35]

Using survey evidence from Chinese urban households, Jin, Li, and Wu (2011) find that greater income inequality is associated with lower consumption and with greater investment in education, where income inequality is measured within age groups by province. Similarly, using high geographical resolution 2001-12 data, Coibion et al. (2014) provide strong evidence that low-income households in high-inequality U.S. locations accumulated less debt (relative to income) than their counterparts in lower-inequality locations. These findings are consistent with Proposition 9, in contrast with the idea that greater information asymmetry about wealth increases wealth-signaling via consumption, or with the intuitive idea that low income individuals borrow and consume more in order to try to keep up with high income households.

Several studies report that wealth dispersion has increased in the United States since the 1980s (e.g. Card and DiNardo (2002), Piketty and Saez (2003), Lemieux (2006)). Given an increase in wealth dispersion, all else equal, Proposition 9 counterfactually implies a rising savings rate. However, the time series shift in *information asymmetry* about others' wealth may have been downward rather than upward, potentially reversing the implication. Our result that wealth dispersion reduces overconsumption derives from asymmetric information—the unobservability of others' wealths—rather than dispersion per se. The rise of the internet has likely made observation of others' wealths or incomes easier than in the past in some countries through search of government or other archives. To the extent that this is true, this effect reinforces the other time series shifts we've described, implying a shift over time toward greater overconsumption.

Furthermore all else was not equal in this time series shift. From the standpoint of our model, a more fundamental effect (which holds even in the base model) comes from the dramatic

reference group income is associated with significantly lower White visible spending. On the other hand, greater dispersion of reference group income is associated with higher visible spending for minorities.

[34]A key parameter in wealth signaling models is the lower support of wealth, which acts as a starting point for the signaling schedule. For any class of bounded wealth distributions that are symmetric with given mean, and such that higher dispersion is associated with more extreme wealth outcomes, it is equivalent to express a result in terms of wealth dispersion or lower support.

[35]Concern for relative consumption, as in 'keeping up with the Joneses' preferences can also induce a fear of falling behind which raises precautionary savings (Harbaugh 1996). Other than Bertrand and Morse (2016), we are not aware of any results in the keeping-up-with-Joneses approach relating overconsumption to wealth dispersion (holding constant the average level of wealth).

transformation of electronic communications and social networks. As discussed in Section 3.1, this has increased the visibility of the consumption activities of others (both absolutely, and relative to non-consumption), which implies greater overconsumption in our model. Also as discussed earlier, the trend toward rising population density, and associated increase in social observation intensity, further reinforces the implication of stronger overconsumption in the base model.

Although overconsumption is decreasing with wealth variance, two considerations are likely to weaken this effect in practice. First, people draw inference from people they regard as peers, who are often in a similar wealth and social class, so that relevant wealth variance may be modest. Second, it is not sheer wealth variation that counts, it is information asymmetry about wealth. Much of real-world variation in wealth is known to others—most people know they are not as rich as Bill Gates or Mark Zuckerberg.

4.3 The Equilibrium Interest Rate

In the base model, the riskfree rate is exogenously set to zero. This corresponds to having storable consumption or, equivalently, to having riskfree bonds in perfectly elastic supply offered at a zero interest rate. We now modify the model to allow for endogenous determination of the interest rate.

When the interest rate can vary, potentially a high interest rate could imply negative date 0 consumption. That does not correspond well with the idea that at worst all consumption bins are empty. We therefore adjust the model to preclude this possibility.

As in the base model, agent utility is defined by (1), where we focus on the case $\rho = 1/2$. Given a one-period interest rate of r, an agent's budget constraint is now

$$c_1 = (1 + r)(W - c_0) - \epsilon, \tag{30}$$

For tractability, in this section we assume a less severe possible wealth disaster than in the base model to preclude negative date 0 consumption over a range of potential interest rates. We therefore assume that

$$\epsilon = \begin{cases} 0 & \text{with probability } \; p \\ \frac{W}{2} & \text{with probability } \; 1 - p, \end{cases} \tag{31}$$

and without loss of generality we focus on the case $W = 1$.

Solving for the optimum of an agent whose probability estimate for a high outcome is $\hat{p}$ yields consumption

$$\frac{1 - r + 2r^2 + (1+r)\hat{p}}{4 + 4r + 2r^2} \stackrel{\text{def}}{=} h + j\hat{p}, \tag{32}$$

with g and f defined in the obvious way. With average agent probability estimate of $\bar{p}$, aggregate per capita consumption is then

$$\bar{c}_0 = h + j\bar{p}. \tag{33}$$

We assume that there is an equally large number of agents who are not subject to ϵ risk, i.e., for whom $c_1 = (1+r)(1-c_0)$. Since these outsiders face no disaster risk, their consumption does not depend on their inferences about p. We can think of them as outsiders such as institutional investors or foreign lenders that supply capital to the individual investors that our analysis focuses upon. Their role in the model is to increase trading opportunities in our exchange economy, so that optimistic beliefs of the individual investors can increase equilibrium per capita consumption, rather than just the interest rate.[36]

Institutional investors are willing to lend to (or borrow from) the agents that are the main focus of our analysis. Since they have no disaster risk, the optimal date 0 consumption of institutions, given r, is

$$\bar{c}_0^I = \frac{1 + r^2}{2 + 2r + r^2}. \tag{34}$$

We assume free disposal of the consumption good, so the equilibrium interest rate satisfies $r \geq -1$. We focus on the region of interest rates in which the institutional investors' lending increases in the interest rate, and therefore require that $r \leq 1/2$.[37]

The per capita endowment of the consumption good at date 0 is fixed, $\bar{c}^e$. Specifically, we assume that the total endowment is such that in an equilibrium with unbiased beliefs (i.e., $\bar{p} = p$), the market clears at interest rate $r = 0$. The market clearing condition is

$$\bar{c}^e = \bar{c}_0^I + \bar{c}_0, \tag{35}$$

where the terms on the right are functions of r. By (33) and (34), and by our assumption that

[36] The consumption of outsiders are excluded from our measure of aggregate consumption; their sole role is to supply capital as an increasing function of the interest rate. Including outsiders in the model is a simplified way of reflecting the idea that in general when the current consumption good is scarce, more can be generated via an aggregate production function for transformation between current and future consumption.

[37] For $r > 1/2$, it is easy to verify that lending decreases in the interest rate. This comes from the standard result in intertemporal choice that an increase in the interest rate has both a substitution effect (which encourages lending) and a wealth effect (which can discourage lending). To illustrate basic insights simply, we focus on the case in which the substitution effect, which is highly intuitive, dominates.

when $r = 0$ a market with unbiased agents would clear, we have

$$\bar{c}^e = \frac{1}{2} + \frac{1+p}{4}, \tag{36}$$

so (35) becomes

$$\bar{c}_0^I + \bar{c}_0 = \frac{1}{2} + \frac{1+p}{4}. \tag{37}$$

Regardless of the level of visibility bias, market clearing implies that r adjust so that aggregate demand is equal to the endowment, so in equilibrium

$$\bar{p} = \frac{(8-5r)r + p(2+2r+r^2)}{2(1+r)}. \tag{38}$$

It is easy to verify that $\bar{p}$ is strictly increasing in r in the relevant region of r.

Along the lines of the arguments in the base model, given r and $\bar{p}$, the fraction of bins that are full is given by (33). Since agents suffer from visibility bias, the fraction of bins *observed* to be full is $S_\tau(h + j\bar{p})$. By Bayes' rule, the agents then arrive at the posterior probability estimate

$$\hat{p} = R(p, S_\tau(h + j\bar{p}), m, h, j), \tag{39}$$

given that the fraction of bins that agents observe are full is z, where the function R is defined in equation (59) in the appendix (see the proof of Proposition 10). An equilibrium is then an outcome in which markets clear, so that (33) holds, and agents' posterior beliefs are in line with their biased observations, $\bar{p} = R(p, S_\tau(h+j\bar{p}), m, h, j)$. For tractability, we focus on the case when $m = 1$ (so that agents put comparable weight on prior information and on social observations).

We show the existence of equilibrium with the following properties:

Proposition 10 *Under the above assumptions, the equilibrium probability estimate, overconsumption, and the interest rate are all increasing in visibility bias, τ.*

5 Concluding Remarks

We examine how bias in social learning endogenously shapes how people trade off current versus future consumption. In our model, people observe the consumption activities of others and use this to update beliefs about whether there is a high or low need to save for the future. Consumption is more salient than non-consumption, resulting in greater observation and cognitive encoding of others' consumption activities. This visibility bias makes episodes of high consumption by others more salient and easier to retrieve from memory than episodes of low consumption. So owing to neglect of selection bias (or a well-known manifestation of it, the availability heuristic), people infer that low saving is warranted. This effect is self-reinforcing at the social level, resulting in overconsumption and high interest rates. With many opportunities to observe others, this feedback effect can be intense. In a social network setting, this effect is further amplified to the extent that in-degree (how heavily an agent is influenced by others) is positively associated with out-degree (how heavily an agent is observed by others) across agents.

The effects in the model can also bring about biased assessments of the savings rates of others, wherein people think that others are consuming even more heavily than they really are. In consequence, a distinctive implication of our approach is that accurate disclosure of the beliefs or average consumption of others can, in some cases, help remedy overconsumption. This implication of our model has been tested in the field experiment of D'Acunto, Rossi, and Weber (2020), who find that, disclosure of average spending causes a greater reduction in the spending of high-spending individuals than the increase in spending by low-spending individuals. The model also implies that not all types of disclosure are always effective in encouraging saving. So to be effective, disclosure policy must be attuned to the geometry of the social network and the demographic structure of the population. The model further predicts that accurate disclosures that increase the salience of *saving* behavior can also reduce overconsumption.

The visibility bias approach offers a simple explanation for a central puzzle in household finance: the dramatic drop in personal saving rates in the U.S. and many other OECD countries over the last 30 years. In the model, greater observability of the consumption of others intensifies the effects of visibility bias, and therefore increases overconsumption. We argue that over the last thirty years the decline in costs of long-distance telephony, the rise of cell phones, cable television and urbanization, and subsequently the rise of the internet, dramatically increased the extent to which people observe possible personal consumption activities of others by television enactment, phone, email, blogging, and social networking. Specifically, this communication is biased toward making the decision to engage rather than not engage in such activities more salient to others, because travel, dining out, or buying a car tend to be relatively noteworthy to report upon.

The visibility bias approach builds upon different elements than other theories of over- or underconsumption. It is inherently social, which distinguishes it from the present bias (hyperbolic discounting) and speculative disagreement theories. Empirically this suggests testing using proxies for sociability, individual network position, and network connectedness, and by population-level characteristics such as wealth variance and age distribution. Some of its predictions are in the opposite direction from other social approaches, such as the wealth signaling (Veblen) and utility interaction (keeping-up-with-the-Joneses) approaches. It also differs from the signaling, preference-based, and speculative disagreement approaches in implying that relatively simple disclosure policy interventions can potentially increase saving. Indeed, there is evidence discussed earlier supporting this implication in specialized settings, such as the decisions of college students of how much to drink.

The model in this paper is static. An interesting extension would be to consider a dynamic setting in which agents consume over time and update in response to common shocks. The feedback/multiplier effect from social learning may then potentially lead to cyclical shifts in overconsumption. Such fluctuations may help explain consumption booms and business cycles. This might potentially provide an interesting contrast to Keynesian ideas about business cycles deriving from resource underutilitization and underconsumption.

Appendices

A Proofs

Proof of Proposition 1: We first describe the model in detail, since some of these were omitted in the main paper.

There are $N = \Theta \times T$ agents in the economy, where $\Theta >> 0$, $T >> 0$, divided into $G \geq 1$ types (types). We will subsequently study a sequence of economies, letting Θ and T tend to infinity at the same rate.

The number of agents of type $g = 1, \ldots, G$ is $N_g >> 0$, and thus

$$N = \sum_{g=1}^{G} N_g.$$

The fraction of agents of type g is $u_g = N_g/N$. Time is discrete, $t = 1, 2, \ldots, T$. Here, a period represents a short time-frame, such as a day or shorter. The T dates together correspond to period 0 as studied in the main paper.

The total potential consumption of an agent is divided into K different activities which we call "bins" (K is large), where each bin represents potential consumption of $W/(2K)$. There are thus in total NK agent-consumption bins. We refer to a bin as *full* if it contains consumption and *empty* otherwise.

In each period, t, Θ of the N agents are selected, and each observes L private unbiased, i.i.d. signals about the level of p. Here, for notational simplicity, we focus on the case when $L_g = L$ for all g. The derivation for the case with heterogeneous L's is identical. Specifically, agent n observes the L Bernoulli distributed signals $\tilde{x}_k^{n,t} \sim \text{Ber}(p)$, $k = 1, \ldots, L$. The agents are chosen in proportion to their type, so the number of selected agents of type g is $u_g\Theta$ (which we assume is an integer for each g) in each period. This is the only information the agents chosen at time 1 receive, whereas agents in later periods also receive signals about the consumptions of other agents in the previous period. We sometimes use the notation $\tilde{x}_k^{n,t,\Theta}$ for the signal $\tilde{x}_k^{n,t}$, i.e., with an extra superscripted index that shows the size of the economy, Θ. This notation, which we use in general for the variables we introduce, will be helpful when taking the limit as the size of the economy grows.

Agents are Bayesians, with prior $\text{Beta}(0, 0)$ distributions about p. After observing the L signals, an agent selected at time 1 therefore updates to the posterior distribution $p \sim \text{Beta}(L^{n,1}, L - L^{n,1})$, where $L^{n,t} = \sum_{k=1}^{L} \tilde{x}_k^{n,t}$. The agent's posterior estimate of the probability of a high outcome is then $\bar{x}^{n,1} = \frac{L^{n,1}}{L}$. Since $E[\tilde{x}_k^{n,1}] = p$, an agent's expected estimate is unbiased, $E[\bar{x}^{n,1}] = p$, and since the number of agents is large, the average estimate across agents selected at $t = 1$ will be very close to p. Moreover, since an agent's consumption is a linear function of p, aggregate consumption will be very close to what is optimal for time 1 agents.

In period 2 and forward, the Θ selected agents observe a sample of consumption bins from the consumption of time $t-1$ agents, in addition to the L unbiased signals. Specifically, each selected agent of type g observes M_g such consumption bins. Here, M_g could for example represent how connected an agent is in a social network (corresponding to the agent's in-degree in the network), a higher connectedness allowing more observations of bins from other agents.

The fraction of the M_g bins an agent of type g observes, that is selected from type g' agents is $u_{g'} > 0$, where $\sum_{g'=1}^{G} u_{g'} = 1$. Here, it may be that $u_{g'} = f_{g'}$ for all g', in which case observations are selected in proportion to the number agents of each type. However, we also allow $u_{g'}$ to be distinct from $f_{g'}$ in general, so that some agent types being more influential (corresponding to them having a higher out-degree in a social network) than the prevalence of their type suggests. So, for example, when $\frac{u_{g'}}{f_{g'}} = 2$, a bin from an agent of type g' is twice as likely to be chosen as would be suggested by the number of agents of that type.

Agents treat these M_g observations in the same way as the L unbiased observations, as being Ber(p) distributed. Indeed, as we shall see, this behavior is rational when full bins are selected with the same probability as their fraction of all bins.

The posterior belief of an agent n of type g at time $t \geq 2$ then is $p \sim Beta(L^{n,t} + M^{n,t}, L + M_g - L^{n,t} - M^{n,t})$, where $M^{n,t}$ is the agent's number of observed full bins, and their probability estimate is

$$\hat{p}^{n,t} = \frac{L^{n,t} + L^{n,t}}{L + M_g} = \frac{\frac{L^{n,t}}{L} + \frac{M_g}{L} \times \frac{M^{n,t}}{M_g}}{1 + \frac{M_g}{L}} = \frac{\bar{x}^{n,t} + m_g \bar{y}^{n,t}}{1 + m_g}, \tag{40}$$

where $m_g = \frac{M_g}{L}$, and $\bar{y}^{n,t} = \frac{M^{n,t}}{M_g}$ is the fraction of observed bins containing consumption. Note that time 1 agents' estimates are consistent with (40) for the special case when $M_g = 0$ (and $m_g = 0$), representing the fact that time 1 agents make no bin observations.

Suppose the fraction $\bar{p}_g^t$ of the bins of agents of type g selected at time t are full. Equivalently, $\bar{p}_g^t$ is the average probability estimate of agents of type g selected at time t. Given that bins from highly connected agents are overrepresented, the fraction of full bins among those considered for selected $t+1$ agents to observe is then

$$\bar{r}^t = \sum_g u_g \bar{p}_g^t,$$

i.e., if B_F^t of these bins are full and B_E^t are empty, then $\bar{r}^t = \frac{B_F^t}{B^t}$, where $B^t = B_F^t + B_E^t$.

In line with our discussion in the main paper, agents are more likely to observe full bins than empty. Specifically, the chance that an *observed* bin is full is

$$\frac{k^F B_F^t}{k^F B_F^t + k^E B_E^t} = \frac{\frac{B_E^t}{B^t}}{\frac{B_F^t}{B^t} + \frac{k^E}{k^F}\left(1 - \frac{B_F^t}{B^t}\right)} = \frac{\bar{r}^t}{\bar{r}^t + \frac{1-\bar{r}^t}{\tau}} \stackrel{\text{def}}{=} S_\tau(\bar{r}^t), \tag{41}$$

where k^F is the probability that a bin is observed conditional upon it being full, k^E is the probability that a bin is observed conditional upon it being empty, and $\tau = k^F/k^E \geq 1$. The parameter τ measures the overrepresentation of full bins in an observer's sample, i.e., it provides a formal definition *visibility bias.* In the benchmark case of $\tau = 1$, the random observations match the actual distribution of consumption bins, and there is no visibility bias. When $\tau > 1$, there is overrepresentation of draws of consumption bins over non-consumption bins.

The average probability estimate of all agents in period t is

$$\bar{p}^{t,\Theta} = \frac{1}{\Theta}\sum_{n=1}^{\Theta} \hat{p}^{n,t} = \sum_{g=1}^{G} f_g \bar{p}_g^t, \tag{42}$$

and the average consumption is $\bar{C}^{t,\Theta} = \bar{p}^{t,\Theta}\frac{W}{2}$. For the growing sequence of economies, we define

$$\bar{p}^t = \lim_{\Theta\to\infty} \bar{p}^{t,\Theta}, \tag{43}$$

which represents the average probability estimate at time t in the large economy. Whereas $\bar{p}^{t,\Theta}$ is random for any finite Θ, we we will show that $\bar{p}^t$ is almost surely a well-defined deterministic number, so aggregate consumption is deterministic in the large economy at all times. We are especially interested in steady-state consumption, after the initial effects of time 1 agents not observing a previous cohort has vanished. We therefore define the average probability estimate among agents in the long-term,

$$\bar{p} = \lim_{t\to\infty} \bar{p}^t. \tag{44}$$

We are now in a position to prove the proposition,. We first study aggregate probability estimates when random realizations are equal to their expected values and show that the result holds in this case. It is then quite straightforward to show that the result holds almost surely for general random realizations of random variables.

Thus, we study $\bar{p}$ when $\bar{x}^{n,t} = E[\bar{x}^{n,t}]$, and $\bar{y}^{n,t} = E[\bar{y}^{n,t}]$ for all n and t, a case we denote by the *expected realization.*

Under the expected realization, it follows immediately that $\bar{p}^{1,g} = p$ for all g and therefore that $\bar{r}^1 = p$, regardless of size, Θ. An iterative application of (40) leads to the relations

$$\bar{p}_g^{t+1} = \frac{p + m_g S_\tau(\bar{r}^t)}{1 + m_g}, \tag{45}$$

$$\bar{r}^{t+1} = \sum_g u_g \bar{p}_g^{t+1}, \tag{46}$$

for all t and g.

We focus on the case $\tau > 1$. From (45,46), it follows that

$$\bar{r}^{t+1} = \sum_g u_g \bar{p}_g^{t+1} = \alpha p + (1 - \alpha) S_\tau(\bar{r}^t) \stackrel{\text{def}}{=} F(\bar{r}^t),$$

regardless of size, Θ, where α is defined in the theorem.

It is easy to verify that $S_\tau(\bar{r})$ is a strictly increasing and concave function in the unit interval, and that $S_\tau(0) = 0$, $S_\tau(1) = 1$. It follows that the function F is also strictly increasing and concave on the unit interval, and that $F(0) > 0$, $F(1) < 1$. This properties of F imply that (16) has a unique root, $\bar{r}$, in the unit interval. A standard fixed point argument implies that all r^t lie within the unit interval and that the sequence $\bar{r}^1, \bar{r}^2, \ldots$ converges to $\bar{r}$ regardless of starting point, $\bar{r}^1 \in (0, 1)$, i.e.,

$$\lim_{t \to \infty} \bar{r}^t = \bar{r}.$$

From (45), it follows that

$$\lim_{t \to \infty} \bar{p}_g^t = \frac{p + m_g S_\tau(\bar{r})}{1 + m_g}, \qquad g = 1, \ldots, G,$$

which leads to

$$\bar{p} = \lim_{t \to \infty} \bar{p}^t = \lim_{t \to \infty} \sum_{g=1}^G f_g \bar{p}_g^t = \sum_{g=1}^G f_g \bar{p}_g.$$

It follows immediately that $\bar{c}_0 = \bar{p}\frac{W}{2}$.

For the special case $\tau = 1$, $S_\tau(\bar{r}^t) = \bar{r}^t$. Now, $\bar{r}^1 = p$, so an iterative application of (45,46) implies that $\bar{p}_g^t = p$ and $\bar{r}^t = p$ for all t. This is in line with $\bar{r} = p$, which is indeed the solution to (16) when $\tau = 1$.

We now extend the proof to allow for random realizations of observed consumption bins. The result follows quite easily from an iterative application of the following standard lemma:

Lemma A.1 *Assume $z_i^s \sim Ber(p^s)$, $i = 1, 2, \ldots, s$, are independent random variables, where $p^s \in [0, 1]$ is a sequence of numbers that converges to $p \in [0, 1]$, $s = 1, 2, \ldots$. Then, almost surely,*

$$\lim_{s \to \infty} \frac{1}{s} \sum_{i=1}^s z_i^s = p.$$

Proof: The result follows immediately from the strong law of large numbers for triangular arrays, since such a set of Bernoulli variables immediately satisfies the fourth moment requirement that $\sup_s E|z_1^s - p^s|^4 \leq 1$. ∎

For the economy of size Θ, consider the agents of type g selected at time $t \leq \Theta$, who update according to the rule (40):

$$\hat{p}^{n,t,\Theta} = \frac{\bar{x}^{n,t,\Theta}}{1+m_g} + \frac{m_g \bar{y}^{n,t,\Theta}}{1+m_g}. \tag{47}$$

Here, and subsequently, we have added the Θ superscript to the notation to keep track of the economy's size, $\bar{x}^{n,t,\Theta} = \frac{1}{\Theta}\sum_{k=1}^{\Theta} \tilde{x}_k^{n,t,\Theta}$, where $\tilde{x}_k^{n,t,\Theta} \sim \text{Ber}(p)$, and $\bar{y}^{n,t,\Theta} = \frac{1}{M_g}\sum_{k=1}^{M_g} \tilde{y}_k^{n,t,\Theta}$, where $\tilde{y}_k^{n,t,\Theta} \sim \text{Ber}(S_\tau(\bar{r}^{t-1,\Theta}))$. The average probability estimate of type g agents at time t is then

$$\bar{p}_g^{t,\Theta} = \frac{1}{\Theta}\sum_{n=1}^{\Theta}\left(\frac{\bar{x}^{n,t,\Theta}}{1+m_g} + \frac{m_g \bar{y}^{n,t,\Theta}}{1+m_g}\right).$$

Now, assume that $\bar{r}^{t-1,\Theta}$ almost surely converges to the number $\bar{r}^{t-1}$, as $\Theta \to \infty$. It then follows from Lemma A.1 that $\bar{p}_g^{t,\Theta}$ converges almost surely to

$$\bar{p}_g^t = \frac{p + m_g S_\tau(\bar{r}^{t-1})}{1+m_g},$$

and thus that $\bar{r}^{t,\Theta}$ converges almost surely to

$$\bar{r}^t = \sum_{g=1}^{G} f_g \bar{p}_g^t.$$

This previous argument also applies to the initial time period, $t = 1$, but with $m_1 = m_2 = \cdots = m_g = 0$, immediately implying that $\bar{p}_g^1 = p$, $g = 1, \ldots, G$, a.s., so $\bar{r}^1 = p$, almost surely. An iterative application of this argument therefore implies that for each t, $\bar{r}^{t,\Theta}$ converges almost surely to the number $\bar{r}^t$ as defined under the expected realization, when Θ tends to infinity, and thus that $\bar{p}_g^t$ converges almost surely to its value defined by (45,46) for each t. Finally, as a consequence, $\bar{p}$, as defined by (42-44), converges almost surely to its value defined by (14-16). We are done.

∎

Proof of Proposition 2: The result follows directly from (15), the fact that $S_\tau(\bar{r}) > p$, and that m_g is increasing in g. ∎

Proof of Proposition 3: Part 1 follows from noting that (19) implies that $\bar{p} = p$ if and only if $-4(1-p)p(\tau-1)^2 m(1+m) = 0$, which holds if and only if $\tau = 1$. Now that $\frac{\partial \bar{p}}{\partial \tau} > 0$ can be seen by substituting $x = \frac{1}{\tau - 1}$, noting that x is decreasing in τ, and taking the derivative w.r.t. x, leading to

$\frac{\partial \bar{p}}{\partial x} = \left(x + p - m + 2pm - \sqrt{(p-x+m)^2 + 4px(1+m)}\right)\psi(x)$, where $\psi(x) > 0$. It then follows from the fact that $(x+p-m+2pm)^2 - ((p-x+m)^2 + 4px(1+m)) = -4(1-p)p\xi(1+m) < 0$, that $\frac{\partial \bar{p}}{\partial x} < 0$, and thus $\frac{\partial \bar{p}}{\partial \tau} > 0$. Part 2 and the claim in 3 that $\bar{p}$ approaches 1 as m becomes large follow immediately by taking the limit of (19) as τ and m become large. To prove the other claims in Part 3, note that $\bar{p}$ can be written as

$$\bar{p} = \frac{\gamma + \sqrt{\gamma^2 + 4(k+\gamma)p}}{2(k+\gamma)} \overset{\text{def}}{=} V(\gamma),$$

where $\gamma = (p+m)(\tau-1) - 1 > -1$, and $k = (1-p)(\tau-1) + 1 > 1$. Since $\frac{\partial \gamma}{\partial m} > 0$, it is therefore sufficient to show that $V'(\gamma) > 0$ when $\gamma > -1$. By calculating $V'(\gamma)$, it follows that $-2\gamma p + k(\gamma - 2p + \sqrt{\gamma^2 + 4p(k+\gamma)}) > 0$ is necessary and sufficient for $V'(\gamma) > 0$ to hold. For $\gamma = 0$, the expression evaluates to $V'(0) = k(-2p + 2\sqrt{kp}) > 0$. Moreover, the solution to $V'(\gamma) = -2\gamma p + k(\gamma - 2p + \sqrt{\gamma^2 + 4kp + 4\gamma p}) = 0$ is $\gamma_{+/-} = -k < -1$. Thus, since V' is a continuous function of γ, $V'(\gamma) > 0$ for all $\gamma \geq -1$. We also verify that the equilibrium probability estimate and aggregate consumption are increasing in the true probability of the high state, p, by calculating $\frac{\partial \bar{p}}{\partial p} = \frac{1}{2(1+m)} + \frac{1}{2\sqrt{V}}\left(2 + \frac{1}{1+m}((\tau-1)(p+m)-1)\right)$, which is obviously positive for $p \in [0,1]$, since $2 + \frac{1}{1+m}((\tau-1)(p+m)-1) \geq 2 + \frac{1}{1+m}((\tau-1)m-1) = 2 + \frac{\tau m}{1+m} - 1 > 0$ for such p. ∎

Proof of Proposition 4: Part 1: The result for $\bar{p}_g$ follows immediately from (15), by noting that the expression represents a weighted average of p and $S_\tau(\bar{r})$ with weight m_g on the second term, and that $S_\tau(\bar{r}) > p$, so when m_g increases, $\bar{p}_g$ indeed increases. The second result then follows immediately from the fact that $\bar{p} = \sum_g f_g \bar{p}_g$ is increasing in each $\bar{p}_g$.

Part 2: The definition of first order stochastic dominance, the fact that $\bar{p}_g$ is increasing in g, and that $\bar{p} = \sum_g f_g \bar{p}_g$ together imply the result.

Part 3: Note that $\frac{f_g^i}{1+m_g}\omega_g^i = \frac{u_g^i}{1+m_g}$, $i \in \{A, B\}$, and thus from (16) that $\alpha^i = \sum_{g=1}^{G} \frac{u_g^i}{1+m_g}$, correspond to the α-coefficient defined in (16). It is straightforward to show that $\bar{r}$ as defined in (16) is decreasing in the α-coefficient defined in (16), and therefore that $\bar{r}^A > \bar{r}^B$ when $\alpha^A < \alpha^B$. Consequently, $\bar{p}_g^A > \bar{p}_g^B$ for all g (as follows from (15)), and $\bar{p}^A > \bar{p}^B$ (as follows from (14)). ∎.

References

Abel, A. B. (1990). Asset prices under habit formation and catching up with the Joneses. *American Economic Review 80*, 38–42.

Akerlof, G. A. and R. J. Shiller (2009). *Animal Spirits: How Human Psychology Drives the Economy and Why it Matters for Global Capitalism*. Princeton, NJ: Princeton University Press.

Allen, T. W. and C. D. Carroll (2001, April). Individual learning about consumption. *Macroeconomic Dynamics 5*(2), 255–271.

Bagwell, L. S. and B. D. Bernheim (1996, June). Veblen effects in a theory of conspicuous consumption. *American Economic Review 86*(3), 349–373.

Bailey, M., R. Cao, T. Kuchler, and J. Stroebel (2018, December). The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy 126*(6), 2224–2276.

Bailey, M., E. Dávila, T. Kuchler, and J. Stroebel (2019, November). House price beliefs and mortgage leverage choice. *Review of Economic Studies 86*(6), 2403–2452.

Benhassine, N., F. Devoto, E. Duflo, P. Dupas, and V. Pouliquen (2015, August). Turning a shove into a nudge? A "labeled cash transfer" for education. *American Economic Journal: Economic Policy 7*(3), 86–125.

Berger, J. and K. L. Milkman (2012, April). What makes online content viral? *Journal of Marketing Research 49*(2), 192–205.

Bertrand, M. and A. Morse (2016, December). Trickle-down consumption. *Review of Economics and Statistics 98*(5), 863–879.

Beshears, J., J. J. Choi, D. Laibson, and B. C. Madrian (2018). Behavioral household finance. In B. D. Bernheim, S. DellaVigna, and D. Laibson (Eds.), *Handbook of Behavioral Economics: Foundations and Applications 1*, Chapter 3, pp. 177–276. Amsterdam, Netherlands: Elsevier.

Beshears, J. L., J. J. Choi, D. Laibson, and B. C. Madrian (2008). The importance of default options for retirement saving outcomes: Evidence from the USA. In S. J. Kay and T. Sinha (Eds.), *Lessons from Pension Reform in the Americas*, Chapter 3, pp. 59–87. Oxford: Oxford University Press.

Beshears, J. L., J. J. Choi, D. Laibson, B. C. Madrian, and K. L. Milkman (2015, June). The effect of providing peer information on retirement savings decisions. *Journal of Finance 70*(3), 1161–1201.

Borgatti, S. P., A. Mehra, D. J. Brass, and G. Labianca (2009, February). Network analysis in the social sciences. *Science 323*(5916), 892–895.

Borgida, E. and R. E. Nisbett (1977, September). The differential impact of abstract vs. concrete information on decisions. *Journal of Applied Social Psychology 7*(3), 258–271.

Bosworth, B. P. (1993, March). *Saving and Investment in a Global Economy*. Washington, D.C.: Brookings Institution Press.

Brenner, L. A., D. J. Koehler, and A. Tversky (1996, March). On the evaluation of one-sided evidence. *Journal of Behavioral Decision Making 9*(1), 59–70.

Brown, J. R., Z. Ivković, P. A. Smith, and S. Weisbenner (2008, June). Neighbors matter: Causal community effects and stock market participation. *Journal of Finance 63*(3), 1509–1531.

Burnside, C., M. Eichenbaum, and S. T. Rebelo (2016, August). Understanding booms and busts in housing markets. *Journal of Political Economy 124*(4), 1088–1147.

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2020, October). Misperceived social norms: Female labor force participation in Saudi Arabia. *American Economic Review 110*(10), 2997–3029.

Cai, H., Y. Chen, and H. Fang (2009, June). Observational learning: Evidence from a randomized natural field experiment. *American Economic Review 99*(3), 864–882.

Campbell, J. Y. and J. H. Cochrane (1999, April). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy 107*(2), 205–251.

Card, D. and J. E. DiNardo (2002, October). Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics 20*(4), 733–783.

Carroll, C. D., B.-K. Rhee, and C. Rhee (1994, August). Are there cultural effects on saving? Some cross-sectional evidence. *Quarterly Journal of Economics 109*(3), 685–699.

Carroll, C. D., B.-K. Rhee, and C. Rhee (1999, October). Does cultural origin affect saving behavior? Evidence from immigrants. *Economic Development and Cultural Change 48*(1), 33–50.

Charles, K. K., E. Hurst, and N. Roussanov (2009, May). Conspicuous consumption and race. *Quarterly Journal of Economics 124*(2), 425–467.

Chen, H., M. Michaux, and N. Roussanov (2019, October). Houses as ATMs? Mortgage refinancing and macroeconomic uncertainty. *Journal of Finance 75*(1), 323–375.

Chen, Y., F. M. Harper, J. Konstan, and S. X. Li (2010, September). Social comparisons and contributions to online communities: A field experiment on movielens. *American Economic Review 100*(4), 1358–1398.

Cialdini, R., L. Demaine, B. Sagarin, D. Barrett, K. Rhoads, and P. Winter (2006, February). Managing social norms for persuasive impact. *Social Influence 1*(1), 3–15.

Coibion, O., Y. Gorodnichenko, M. Kudlyak, and J. Mondragon (2014, January). Does greater inequality lead to more household borrowing? New evidence from household data. Working paper, University of Texas Austin.

Cole, H. L., G. J. Mailath, and A. Postlewaite (1995, Summer). Incorporating concern for relative wealth into economic models. *Quarterly Review, Federal Reserve Bank of Minneapolis 19*(3), 12–21.

Corneo, G. and O. Jeanne (1997, October). Conspicuous consumption, snobbism and conformism. *Journal of Public Economics 66*(1), 55–71.

D'Acunto, F., A. Rossi, and M. Weber (2020, June). Crowdsourcing peer information to change spending behavior. Working paper, Boston College.

de Zúñiga, H. G. and S. Valenzuela (2011, December). The mediating path to a stronger citizenship: Online and offline networks, weak ties, and civic engagement. *Communication Research 38*(3), 397–421.

DeGroot, M. H. and M. J. Schervish (2012). *Probability and Statistics.* Boston, MA: Pearson.

Donath, J. and D. Boyd (2004, October). Public displays of connection. *BT Technology Journal 22*(4), 71–82.

Duflo, E. and E. Saez (2002, July). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. *Journal of Public Economics 85*(1), 121–148.

Duflo, E. and E. Saez (2003, August). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics 118*(3), 815–842.

Dupor, B. and W.-F. Liu (1993, March). Jealousy and equilibrium overconsumption. *American Economic Review 93*(1), 423–428.

Enke, B. (2020, August). What you see is all there is. *Quarterly Journal of Economics 135*(3), 1363–1398.

Federal Reserve Board (2018, May). Report on the economic well-being of U.S. households in 2017.

Fiedler, K. (2008, January). The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition 34*(1), 186–203.

Frederick, S. (2005, Autumn). Cognitive reflection and decision making. *Journal of Economic Perspectives 19*(4), 25–42.

Frederick, S. (2012, June). Overestimating others' willingness to pay. *Journal of Consumer Research 39*(1), 1–21.

Frey, B. and S. Meier (2004, December). Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. *American Economic Review 94*(5), 1717–1722.

Galí, J. (1994, February). Keeping up with the Joneses: Consumption externalities, portfolio choice, and asset prices. *Journal of Money, Credit and Banking 26*(1), 1–8.

Georgarakos, D., M. Haliassos, and G. Pasini (2014, May). Household debt and social interactions. *Review of Financial Studies 27*(5), 1404–1433.

Georgarakos, D. and G. Pasini (2011, October). Trust, sociability, and stock market participation. *Review of Finance 15*(4), 693–725.

Gerber, A. S. and T. Rogers (2009, January). Descriptive social norms and motivation to vote: Everybody's voting and so should you. *The Journal of Politics 71*(1), 178–191.

Giorgi, G. D., A. Frederiksen, and L. Pistaferri (2016, June). Consumption network effects. NBER Working Paper No. 22357.

Goldstein, N. J., R. B. Cialdini, and V. Griskevicius (2008, October). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research 35*(3), 472–482.

Golub, B. and E. Sadler (2016). Learning in social networks. In Y. Bramoullé, A. Galeotti, and B. W. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, Chapter 19, pp. 504–542. New York, NY: Oxford University Press.

Gomes, F., K. Hoyem, W.-Y. Hu, and E. Ravina (2020, February). Retirement savings adequacy in U.S. defined contribution plans. Working paper, London Business School.

Grinblatt, M., M. Keloharju, and S. Ikäheimo (2008, November). Social influence and consumption: Evidence from the automobile purchases of neighbors. *Review of Economics and Statistics 90*(4), 735–753.

Gross, D. B. and N. S. Souleles (2002, February). Do liquidity constraints and interest rates matter for consumer behavior? Evidence from credit card data. *Quarterly Journal of Economics 117*(1), 149–185.

Guidolin, M. and E. A. L. Jeunesse (2007, November/December). The decline in the U.S. personal saving rate: Is it real and is it a puzzle? *Federal Reserve Bank of St. Louis Review 89*(6), 491–514.

Guin, B. (2017, November). Culture and household saving. Working paper, Bank of England.

Han, B., D. Hirshleifer, and J. Walden (2020, June). Social Transmission Bias and Investor Behavior. *Journal of Financial and Quantitative Analysis*. Forthcoming.

Harbaugh, R. (1996, December). Falling behind the Joneses: relative consumption and the growth-savings paradox. *Economics Letters 53*(3), 297–304.

Healy, A. F. (1981, September). The effects of visual similarity on proofreading for misspellings. *Memory and Cognition 9*(5), 453–460.

Hearst, E. (1991). Psychology and nothing. *American Scientist 79*(5), 432–443.

Heimer, R. (2016, November). Peer pressure: Social interaction and the disposition effect. *Review of Financial Studies 29*(11), 3177–3209.

Heyerdahl-Larsen, C. and J. Walden (2017, April). Welfare in economies with production and heterogeneous beliefs. Working paper, London Business School.

Hirshleifer, D. and S. H. Teoh (2009). Thought and Behavior Contagion in Capital Markets. In T. Hens and K. Schenk-Hoppe (Eds.), *Handbook of Financial Markets: Dynamics and Evolution*, Handbooks in Finance, Chapter 1, pp. 1–46. Amsterdam, The Netherlands: North-Holland.

Hong, H. G., J. D. Kubik, and J. C. Stein (2004, February). Social interaction and stock market participation. *Journal of Finance 59*(1), 137–163.

Jackson, M. O. (2019, April). The friendship paradox and systematic biases in perceptions and social norms. *Journal of Political Economy 127*(2), 777–818.

Jin, Y., H. Li, and B. Wu (2011, June). Income inequality, consumption, and social-status seeking. *Journal of Comparative Economics 39*(2), 191–204.

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux.

Kahneman, D. and A. Tversky (1972, July). Subjective probability: A judgment of representativeness. *Cognitive Psychology 3*(3), 430–454.

Kahneman, D. and A. Tversky (1973, July). On the psychology of prediction. *Psychological Review 80*(4), 237–25l.

Kaustia, M. and S. Knüpfer (2012, May). Peer performance and stock market entry. *Journal of Financial Economics 104*(2), 321–338.

Koehler, J. J. and M. Mercer (2009, July). Selection neglect in mutual fund advertisements. *Management Science 55*(7), 1107–1121.

Krapivsky, P. L., G. J. Rodgers, and S. Redner (2001, June). Degree distributions of growing networks. *Physical Review Letters 86*(23), 5401.

Laibson, D. (1997, May). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics 112*(2), 443–477.

Laibson, D. I., A. Repetto, and J. Tobacman (1998). Self-control and saving for retirement. *Brookings Papers on Economic Activity 1998*(1), 91–196.

Lambert, T. A., A. S. Kahn, and K. J. Apple (2003, May). Pluralistic ignorance and hooking up. *Journal of Sex Research 40*(2), 129–133.

Lemieux, T. (2006, May). Postsecondary education and increasing wage inequality. *American Economic Review 96*(2), 195–199.

Lieber, E. M. and W. Skimmyhorn (2018, July). Peer effects in financial decision-making. *Journal of Public Economics 163*, 37–59.

Ljungqvist, L. and H. Uhlig (2000, June). Tax policy and aggregate demand management under catching up with the Joneses. *American Economic Review 90*(3), 356–366.

Loayza, N., K. Schmidt-Hebbel, and L. Serven (2000, September). Saving in developing countries: An overview. *World Bank Economic Review 14*(3), 393–414.

Lusardi, A., O. S. Mitchell, and N. Oggero (2018, May). The changing face of debt and financial fragility at older ages. *AEA Papers and Proceedings 108*, 407–411.

Madrian, B. and D. Shea (2001, November). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics 116*(4), 1149–1187.

Malmendier, U. and L. S. Shen (2018, June). Scarred consumption. Working paper, NBER Working Paper No. 24696.

McShane, B. B., E. T. Bradlow, and J. Berger (2012, December). Visual influence and social groups. *Journal of Marketing Research 49*(6), 854–871.

Mok, D., B. Wellman, and J. Carrasco (2010, November). Does distance matter in the age of the Internet? *Urban Studies 47*(13), 2747–2783.

Moretti, E. (2011, January). Social learning and peer effects in consumption: Evidence from movie sales. *Review of Economic Studies 78*(1), 356–393.

Neisser, U. (1963, September). Decision-time without reaction-time: Experiments in visual scanning. *American Journal of Psychology 76*(3), 376–385.

Nisbett, R. and L. Ross (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.

OECD (2014, August). OECD factbook 2014: Economic, environmental and social statistics. *OECD Publishing 89*(6).

Parker, J. A. (1999, January). Spendthrift in America? On two decades of decline in the U.S. saving rate. *NBER Macroeconomics Annual 14*, 317–370.

Parlour, C. A., V. Prasnikar, and U. Rajan (2007, August). Compensating for the winner's curse: Experimental evidence. *Games and Economic Behavior 60*(2), 339–356.

Perkins, H. W., M. P. Haines, and R. Rice (2005). Misperceiving the college drinking norm and related problems: A nationwide study of exposure to prevention information, perceived norms and student alcohol misuse. *Journal of Studies on Alcohol 66*(4), 470–478.

Perkins, H. W., P. W. Meilman, J. S. Leichliter, J. R. Cashin, and C. A. Presley (1999). Misperceptions of the norms for the frequency of alcohol and other drug use on college campuses. *Journal of American College Health 47*(6), 253–258.

Piketty, T. and E. Saez (2003, February). Income inequality in the United States, 1913-1998. *Quarterly Journal of Economics 118*(1), 1–39.

Poterba, J. M. (2014, May). Retirement security in an aging population. *American Economic Review 104*(5), 1–30.

Poterba, J. M., S. F. Venti, and D. A. Wise (2012, May). Were they prepared for retirement? Financial status at advanced ages in the HRS and AHEAD cohorts. In D. A. Wise (Ed.), *Investigations in the Economics of Aging*, pp. 21–69. University of Chicago Press.

Prentice, D. A. and D. T. Miller (1993, February). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology 64*(2), 243–256.

Salganik, M. J., P. S. Dodds, and D. J. Watts (2006, February). Experimental study of inequality and unpredictability in an artificial cultural market. *Science 311*(5762), 854–856.

Samuelson, W. and R. Zeckhauser (1988, March). Status-quo bias in decision making. *Journal of Risk and Uncertainty 1*(1), 7–59.

Scellato, S., C. Mascolo, M. Musolesi, and V. Latora (2010, June). Distance matters: Geo-social metrics for online social networks. In *Proceedings of the 3rd Conference on Online Social Networks*, WOSN'10, pp. 8–8.

Scholz, J. K., A. Seshadri, and S. Khitatrakun (2006, August). Are Americans saving "optimally" for retirement? *Journal of Political Economy 114*(4), 607–643.

Schroeder, C. M. and D. A. Prentice (1998, December). Exposing pluralistic ignorance to reduce alcohol use among college students. *Journal of Applied Social Psychology 28*(23), 2150–2180.

Shefrin, H. M. and R. H. Thaler (1988, October). The behavioral life-cycle hypothesis. *Economic Inquiry 26*(4), 609–643.

Shemesh, J. and F. Zapatero (2016, September). The intensity of keeping up with the Joneses behavior: Evidence from neighbor effects in car purchases. Working paper, Monash University.

Stanford Center on Longevity (2016, February). Seeing our way to living long, living well in 21st century America. *The Sightlines Project*.

U.S. Government Accountability Office (2015, May). Most households approaching retirement have low savings.

White, K. J. C. and A. M. Guest (2003, September). Community lost or transformed? Urbanization and social ties. *City & Community 2*(3), 239–259.

Internet Appendix

Proof of Proposition 5: The proof is identical to that of Proposition 1, but with included disclosure signals. ∎

Proof of Proposition 6: From (22) it follows that

$$\begin{aligned}\left(1+m_1+\frac{Q}{L_1}(1-v_1)\right)\bar{p}_1-\frac{Q}{L_1}(1-v_1)\bar{p}_2 &= p+m_1S,\\ \left(1+m_2+\frac{Q}{L_2}v_1\right)\bar{p}_2-\frac{Q}{L_2}v_1\bar{p}_1 &= p+m_2S,\end{aligned}$$

where we write $S=S_\tau(\bar{r})$ for notational convenience. We rewrite this system on matrix form as

$$\mathbf{A}\mathbf{p}=p\mathbf{1}+S\mathbf{m},$$

where

$$\mathbf{p}=\begin{bmatrix}\bar{p}_1\\ \bar{p}_2\end{bmatrix},\qquad \mathbf{A}=\begin{bmatrix}1+m_1+\frac{Q}{L_1}(1-v_1) & -\frac{Q}{L_1}(1-v_1)\\ -\frac{Q}{L_2}v_1 & 1+m_2+\frac{Q}{L_2}v_1\end{bmatrix},\qquad \mathbf{m}=\begin{bmatrix}m_1\\ m_2\end{bmatrix}.$$

Solving for $\mathbf{p}$, we arrive at

$$\mathbf{p}=\mathbf{A}^{-1}(p\mathbf{1}+S\mathbf{m}),\qquad \text{where} \tag{48}$$

$$\mathbf{A}^{-1}=\frac{1}{R}\begin{bmatrix}1+m_2+\frac{Q}{L_2}v_1 & \frac{Q}{L_1}(1-v_1)\\ \frac{Q}{L_2}v_1 & 1+m_1+\frac{Q}{L_1}(1-v_1)\end{bmatrix},\qquad \text{and}$$

$R=(1+m_1)(1+m_2)+(1+m_2)(1-v_1)\frac{Q}{L_1}+(1+m_1)v_1\frac{Q}{L_2}$.

We also have $\bar{p}=(f_1,f_2)\cdot\mathbf{p}$ (representing (22) on matrix form), which via (48) leads to

$$\begin{aligned}\bar{p} &= \frac{1}{R}\Big(f_1(1+m_2)(p+m_1S)+f_2(1+m_1)(p+m_2S)+ \\ &+\ Q\Big(\frac{v_1}{L_2}(p+m_1S)+\frac{1-v_1}{L_1}(p+m_2S)\Big)\Big) \\ &\stackrel{\text{def}}{=} \frac{b+Qa}{d+Qc},\qquad \text{where}\end{aligned} \tag{49}$$

$$\begin{aligned} a &= \frac{v_1}{L_2}(p+m_1S)+\frac{1-v_1}{L_1}(p+m_2S),\\ b &= f_1(1+m_2)(p+m_1S)+f_2(1+m_1)(p+m_2S),\\ c &= \frac{v_1}{L_2}(1+m_1)+\frac{1-v_1}{L_1}(1+m_2)\\ d &= (1+m_1)(1+m_2).\end{aligned}$$

We write (49) as $\bar{p}=\bar{p}(S,Q)$, and the chain rule then implies that

$$\frac{d\bar{p}}{dQ}=\frac{\partial\bar{p}}{\partial Q}+\frac{\partial\bar{p}}{\partial S}\frac{dS}{dQ}. \tag{50}$$

A necessary and sufficient condition for $\frac{\partial \bar{p}}{\partial Q} > 0$ is that $ad - bc > 0$, which one can verify is equivalent to

$$v_1 > \frac{1}{\frac{L_1(1+m_1)}{L_2(1+m_2)}\left(\frac{1}{f_1} - 1\right) + 1}.$$

Similarly, $\frac{\partial \bar{p}}{\partial Q} < 0$ is equivalent to

$$v_1 < \frac{1}{\frac{L_1(1+m_1)}{L_2(1+m_2)}\left(\frac{1}{f_1} - 1\right) + 1}.$$

By inspection, one verifies that $\frac{\partial \bar{p}}{\partial S} > 0$ (note that R does not depend on S). Also, $\frac{dS}{dQ} = \frac{\partial S}{\partial \beta}\frac{d\beta}{dQ}$, where from (23) it follows that $\frac{\partial S}{\partial \beta} < 0$. From (24), moreover, it follows that

$$\beta = \frac{L_1 L_2(1 + m_1 u_2 + m_2 u_1) + Q(L_2(1 - v_1) + L_1 v_1)}{L_1 L_2(1 + m_1)(1 + m_2) + Q(v_1 L_1(1 + m_1) + (1 - v_1)L_2(1 + m_2))} \stackrel{\text{def}}{=} \frac{b' + Qa'}{d' + Qc'}. \tag{51}$$

A necessary and sufficient condition for $\frac{d\beta}{dQ} < 0$ (so that $\frac{dS}{dQ} > 0$) is that $a'd' - b'c' < 0$, which in turn is equivalent to

$$v_1 > \frac{1}{\frac{L_1(1+m_1)}{L_2(1+m_2)}\left(\frac{1}{u_1} - 1\right) + 1}.$$

Similarly, $\frac{d\beta}{dQ} > 0$ (so that $\frac{dS}{dQ} < 0$) is equivalent to

$$v_1 < \frac{1}{\frac{L_1(1+m_1)}{L_2(1+m_2)}\left(\frac{1}{u_1} - 1\right) + 1}.$$

Altogether, these conditions when plugged into (50) imply parts 1 and 2 of the proposition.

To show part 3, rewrite (49) as

$$\bar{p}(S, v_1) = \frac{b_2 + v_1 a_2}{d_2 + v_1 c_2}, \qquad \text{where}$$

$$\begin{aligned}
a_2 &= Q\left(\frac{1}{L_2}(p + m_1 S) - \frac{1}{L_1}(p + m_2 S)\right), \\
b_2 &= f_1(1 + m_2)(p + m_1 S) + f_2(1 + m_1)(p + m_2 S) + \frac{Q}{L_1}(p + m_2 S), \\
c_2 &= Q\left(\frac{1}{L_2}(1 + m_1) - \frac{1}{L_1}(1 + m_2)\right) \\
d_2 &= (1 + m_1)(1 + m_2) + \frac{Q}{L_1}(1 + m_2),
\end{aligned}$$

and

$$\beta \stackrel{\text{def}}{=} \frac{b_3 + v_1 a_3}{d_3 + v_1 c_3}.$$

The chain rule gives us

$$\frac{d\bar{p}}{dv_1} = \frac{\partial \bar{p}}{\partial v_1} + \frac{\partial \bar{p}}{\partial S}\frac{\partial S}{\partial \beta}\frac{d\beta}{dv_1}. \tag{52}$$

where, as noted earlier, $\frac{\partial \bar{p}}{\partial S} > 0$, and $\frac{\partial S}{\partial \beta} < 0$. A similar comparison as before shows that

$$a_3 d_3 - b_3 c_3 = L_1 L_2 (m_2 - m_1) Q (Q + L_1 (1 + m_1) u_2 + L_2 (1 + m_2) u_1) > 0,$$

and thus $\frac{d\beta}{dv_1} > 0$, in following the earlier argument. Moreover, when $L_1 = L_2 = L$, we get

$$a_2 d_2 - b_2 c_2 = -L^2 (m_2 - m_1) Q (Q + L(1 + m_1 u_2 + m_2 u_1) < 0,$$

so $\frac{\partial \bar{p}}{\partial v_1} < 0$. Altogether, using (52), this then implies $\frac{d\bar{p}}{dv_1} < 0$. We are done. ∎

Proof of Corollary 1: 1. and 2. follow immediately from the fact that $0 < \underline{v} < \bar{v} < 1$.

3(a): When $\omega_1 = \omega_2$, $u_1 = f_1$, and moreover per assumption $v_1 = f_1$, so $u_1 = f_1 = v_1$. Moreover, $L_1 = L_2$, so it is sufficient to show that

$$f_1 < \underline{v} = \frac{1}{\frac{1+m_1}{1+m_2}\left(\frac{1}{f_1} - 1\right) - 1},$$

which follows immediately since $m_1 < m_2$.

3(b): Since $v_1 = f_1$, it is enough to show that $f_1 > \bar{v}$, which is equivalent to the stated condition. ∎

Proof of Proposition 7: We study a stationary equilibrium, in which $\bar{p}$ is constant over time. Denote by $\bar{p}$ the equilibrium probability estimate in (25) and by $\bar{p}$' the equilibrium in (15). It follows from these equations that

$$\bar{p}' - \bar{p} = \frac{m}{1+m}(S_\tau(\bar{p}') - S_\tau(\bar{p})) + \frac{q}{1+m}(1 - \bar{p}' - S_\tau(1 - \bar{p}')). \tag{53}$$

Define the function

$$R(x) = x - \bar{p} - \left(\frac{m}{1+m}(S_\tau(x) - S_\tau(\bar{p})) + \frac{q}{1+m}(1 - x - S_\tau(1 - x))\right), \tag{54}$$

and note that $R(\bar{p}') = 0$, and that R is a continuous function of $x \in [0, 1]$. From (17), it follows that $\bar{p} > \frac{m}{1+m} S_\tau(\bar{p})$, and therefore, since $S_\tau(0) = 0$, $S_\tau(1) = 1$, that $R(0) < 0$. Moreover, since $S_\tau(y) > y$, for $y \in (0, 1)$, it follows that $R(\bar{p}) = \frac{q}{1+m}(S_\tau(1 - \bar{p}) - (1 - \bar{p})) > 0$. Thus, by the intermediate value theorem, it follows that the point, $\bar{p}'$ where $R(\bar{p}') = 0$, satisfies $0 < \bar{p}' < \bar{p}$. ∎

Proof of Proposition 8: It is easy to verify that the solution to the equilibrium condition (26) is

$$\begin{aligned} \bar{p} \;&=\; \frac{-1}{2(2\lambda - 1)(1 + m)(\tau - 1)}\Big(1 - 3p + 4\lambda p + m - 2pm + 2\lambda pm + 3p\tau - \\ &-\; 4\lambda p\tau + m\tau + 2pm\tau - 2\lambda pm\tau - \sqrt{V}\Big), \end{aligned} \tag{55}$$

where

$$\begin{aligned} V \;&=\; -4(2\lambda - 1)p(1 + m)(\tau - 1)\Big(-1 + 2(\lambda - 1)p(\tau - 1) + 2(\lambda - 1)m\tau\Big) + \\ &+\; \Big(1 - p(-3 - 2m + 2\lambda(2 + m)(\tau - 1) + m(1 + \tau - 2\lambda\tau)\Big)^2. \end{aligned}$$

It is also easy to verify that the solution reduces to (27,28) when $\lambda = 0$, and to (19) when $\lambda = 1$. Moreover, it is easy to verify that $\bar{p}^{\lambda=0} > p$, and that for any $\lambda \in [0, 1]$, $\bar{p}^\lambda = p \Rightarrow p \in \{0, 1\}$. It also follows immediately that $\bar{p}$ is a continuous function of λ, except possibly at $\lambda = 1/2$.

We next define $x = 2\lambda - 1 \in [-1, 1]$, and rewrite (55) as

$$\begin{aligned}
\bar{p} &= a + \frac{b(\sqrt{1 - cx + dx^2} - 1)}{2x(\tau - 1)}, \qquad \text{where} \\
a &= \frac{p(2+m)(\tau - 1) + m\tau}{2(1+m)(\tau - 1)}, \\
b &= 1 + p(\tau - 1) \\
c &= \frac{2m(-p(\tau-1)^2 + p^2(\tau-1)^2 + \tau)}{(1+m)(1+p(\tau-1))^2}, \\
d &= \frac{m^2(p(1-\tau)+\tau)^2}{(1+m)^2(1+p(\tau-1))^2}.
\end{aligned}$$

A Taylor expansion of $\sqrt{1 + cx + dx^2} - 1$ around $x = 0$ i.e., $\lambda = 1/2$, yields $\sqrt{1 + cx + dx^2} - 1 = \frac{1}{2}cx + O(x^2)$, and thus $\bar{p}$ is a continuous function of λ at $\lambda = 1/2$ too. Thus, since $\bar{p}^{\lambda=1} > p$, $\bar{p}^{\lambda}$ depends continuously on λ, and $\bar{p}^{\lambda} \neq p$ for $\lambda \in [0, 1]$, it follows that $\bar{p}^{\lambda} > p$ for all $\lambda \in [0, 1]$. We have shown $\bar{p} > p$, i.e., (1), and (3).

To show (2), we note that

$$\frac{d\bar{p}}{dx} = \frac{b}{4(\tau - 1)} \times \frac{\sqrt{1 - cx + dx^2} + \frac{c}{2}x - 1}{\sqrt{1 - cx + dx^2}},$$

so $\sqrt{1 - cx + dx^2} > 1 - \frac{c}{2}x$ is necessary and sufficient for $\frac{d\bar{p}}{dx} > 0$. This implies the following sufficient condition:

$$1 - cx + dx^2 > \left(1 - \frac{c}{2}x\right)^2 = 1 - cx + \frac{c^2}{4}x^2,$$

or equivalently,

$$4d^2 - c^2 > 0.$$

It is easy to verify that

$$4d^2 - c^2 = \frac{16(1-p)pm^2(\tau-1)^2\tau}{(1+m)^2(1+p(\tau-1))^2} > 0,$$

so the condition is indeed satisfied. ∎

Proof of Proposition 9:

We study a stationary equilibrium, in which $\bar{p}$ is constant over time. We first state and prove the following lemma, which characterizes the equilibrium probability estimate:

Lemma A.2 *The equilibrium probability estimate is the solution to the equation*

$$\begin{aligned}
\bar{p} &= \lambda R(p, S_\tau(f\bar{p}(1-\Delta)), m, f(1+\Delta)) \\
&+ (1 - 2\lambda) R(p, S_\tau(f\bar{p}), m, f(1+\Delta)) \\
&+ \lambda R(p, S_\tau(f\bar{p}(1+\Delta)), m, f(1+\Delta)), \qquad (56)
\end{aligned}$$

where the function R is defined by

$$R(p, z, m, f) = \frac{1}{2f(1+m)}\Big(1 + fp + fm + zm - \sqrt{(1 + fp + fm + zm)^2 - 4f(1+m)(p+zm)}\Big).$$

The lemma states that an agent who observes higher-than-expected consumption updates beliefs as if the agent were observing only wealthy agents (who consume the fraction $\bar{p}f(1+\Delta)$ of bins rather than the average, $\bar{p}f$). The reason why the agent so strongly concludes that the wealthy were observed is that the number of observations Q and M are large. The observer finds the strength of the evidence of high consumption very surprising; the likelihood is low under the hypothesis that observations are of either high or low wealth agents. But the likelihood is especially low when the observation targets have low wealth, so the posterior belief puts all the weight on observing wealthy agents.

Proof of Lemma A.2: For $\alpha, \beta, f_1, f_2 \in (0,1]$, $m > 0$, define

$$L(\alpha,\beta,m,f_1,f_2) \quad = \quad \lim_{Q\to\infty} \frac{J(\alpha,\beta,m,f_1,Q,0)}{J(\alpha,\beta,m,f_2,Q,0)},$$

where J was previously defined. It follows from standard properties of Beta distributions, that

$$L(\alpha,\beta,m,f_1,f_2) = \begin{cases} \infty, & \left|f_1 - \frac{\beta}{\alpha}\right| < \left|f_2 - \frac{\beta}{\alpha}\right|, \\ \\ 0, & \left|f_1 - \frac{\beta}{\alpha}\right| > \left|f_2 - \frac{\beta}{\alpha}\right|. \end{cases} \tag{57}$$

An agent with prior p, who observes a fraction β of bins with consumption, believing that the observations provide an unbiased estimate of the consumption of others, and who believes that the distribution of wealth groups (poor, medium, rich) among the population is $(\lambda, 1-2\lambda, \lambda)$ who consume the fraction $(f(1-\Delta), f, f(1+\Delta))$ of the bins, respectively, will update—using Bayes rule—to the posterior:

$$\begin{aligned} \hat{p} &= \frac{\lambda J(\alpha,\beta,m,(1-\Delta)f,Q,1) + (1-2\lambda)J(\alpha,\beta,m,f,Q,1) + \lambda J(\alpha,\beta,m,(1-\Delta)f,Q,1)}{\lambda J(\alpha,\beta,m,(1+\Delta)f,Q,0) + (1-2\lambda)J(\alpha,\beta,m,f,Q,0) + \lambda J(\alpha,\beta,m,(1+\Delta)f,Q,0)} \\ &= \frac{J(\alpha,\beta,m,(1-\Delta)f,Q,1)}{J(\alpha,\beta,m,(1-\Delta)f,Q,0)} g_1 + \frac{J(\alpha,\beta,m,f,Q,1)}{J(\alpha,\beta,m,f,Q,0)} g_2 + \frac{J(\alpha,\beta,m,(1+\Delta)f,Q,1)}{J(\alpha,\beta,m,(1+\Delta)f,Q,0)} g_3, \end{aligned}$$

where

$$\begin{aligned} g_1 &= \frac{\lambda J(\alpha,\beta,m,(1-\Delta)f,Q,0)}{\lambda J(\alpha,\beta,m,(1-\Delta)f,Q,0) + (1-2\lambda)J(\alpha,\beta,m,f,Q,0) + \lambda J(\alpha,\beta,m,(1+\Delta)f,Q,0)}, \\ g_2 &= \frac{(1-2\lambda)J(\alpha,\beta,m,f,Q,0)}{\lambda J(\alpha,\beta,m,(1-\Delta)f,Q,0) + (1-2\lambda)J(\alpha,\beta,m,f,Q,0) + \lambda J(\alpha,\beta,m,(1+\Delta)f,Q,0)}, \\ g_3 &= \frac{\lambda J(\alpha,\beta,m,(1+\Delta)f,Q,0)}{\lambda J(\alpha,\beta,m,(1-\Delta)f,Q,0) + (1-2\lambda)J(\alpha,\beta,m,f,Q,0) + \lambda J(\alpha,\beta,m,(1+\Delta)f,Q,0).} \end{aligned}$$

It follows from (57) and assumption (29), that as $Q \to \infty$, $g_1, g_2 \to 0$ and $g_3 \to 1$. In words, the observing agent puts all the weight on having observed a wealthy agent's consumption, regardless of whom the observing agent actually observes. Moreover, from (65) and it follows that an agent observing poor, average, and wealthy agents consuming based on posterior beliefs $\bar{p}$ will have posterior belief

$$\hat{p} = R(\alpha,\beta,m,(1+\Delta)f),$$

where β equals $S_\tau((1-\Delta)f\bar{p})$, $S_\tau(f\bar{p})$, and $S_\tau((1+\Delta)f\bar{p})$ with probabilities λ, $1-2\lambda$, and λ, respectively. The fixed point problem that matches aggregate posterior beliefs with agents' updating is therefore (56).

Existence of a solution to (56) follows from the easily verifiable fact that $R(p, S_\tau(g\times 0), m, f) > 0$ and $R(p, S_\tau(g \times 1), m, f) < 1$, regardless of $g, p, f \in (0,1)$, and $m > 0$. Therefore, the r.h.s of

(56), which is a continuous function, is strictly greater than $\bar{p}$ when $\bar{p}$ close to zero, and strictly less than $\bar{p}$ when $\bar{p}$ is close to one. Existence therefore follows from the intermediate value theorem. This completes the proof of Lemma A.2. ∎

It is straightforward to verify that the function R satisfies $\frac{\partial R}{\partial z} > 0$, since

$$\frac{\partial R}{\partial z} = \frac{m(1+c)}{2f(1+m)},$$

where

$$c^2 = \frac{(-1 - zm + f(2-p+m))^2}{(f^2(p+m)^2 + (1+\beta m)^2 + 2f(m - \beta m(2+m) + \alpha(-1+(\beta-2)m))} < 1,$$

implying positivity of the derivative. It also follows that $\frac{\partial^2 R}{\partial z^2} = -\frac{2(1-f)(1-p)m^2}{((1+fp+fm+zm)^2 - 4f(1+m)(p+zm))^{3/2}} < 0$, so R is concave in z.

Moreover, one can show that $\frac{\partial R}{\partial f} < 0$. Specifically, it is easy to verify that $\frac{\partial R}{\partial f} = \frac{\kappa_1}{\kappa_2 f^2}$, where the function $\kappa_2 > 0$, and $\kappa_1 = 0 \Leftrightarrow f = 0$ on $f \in [0,1]$, for the smooth function κ_1. Thus,R is a monotone function for positive $0 \le f \le 1$. A Taylor expansion of κ_1 in f around $f = 0$ implies that κ_1 is of the form $R'(f) = -c_1 f^2 + O(f^3)$, where the constant $c_1 > 0$, altogether implying that $\frac{\partial R}{\partial f} < 0$ for small positive f, and thereby for all $0 \le f \le 1$ (since R is monotone).

Now, we use these properties of R to show that the total derivative of the r.h.s. of (56) w.r.t. Δ is negative. Specifically, using the notation R_i for the partial derivative of the function R w.r.t. its ith argument, from the calculus of total derivatives it follows that this r.h.s. derivative is of the form

$$q\bar{p}f\left(-S'_\tau(f\bar{p}(1-\Delta))R_2(\cdot) + S'_\tau(f\bar{p}(1+\Delta))R_2(\cdot)\right) + f(qR_4(\cdot) + (1-2q)R_4(\cdot) + qR_4(\cdot)).$$

Since $R_4(\cdot) < 0$, the second part of this expression is negative. Moreover, S_τ is concave and R is concave in its second argument, so the first part of the expression is also negative. Thus, the r.h.s. of (56) is decreasing in Δ.

Because R is increasing and concave in its second argument, it follows that the r.h.s. of (56) is concave and increasing in $\bar{p}$, and since $R(p,0,m,f) > 0$, it follows that at the equilibrium point $0 < \frac{\partial R}{\partial \bar{p}} < 1$. Altogether, the inverse function theorem then implies that $\frac{\partial \bar{p}}{\partial \Delta} < 0$ for the fixed point $\bar{p}$ defined by (56). ∎

Proof of Proposition 10: We study a stationary equilibrium, in which $\bar{p}$ is constant over time. The proof of the Bayesian updating follows similar lines as in Proposition A.2. Define

$$J(\alpha,\beta,m,f,g,Q,x) = \int_0^1 t^{\alpha Q - 1 + x}(1-t)^{(1-\alpha)Q-1}(g+ft)^{\beta m Q}(1-g-ft)^{(1-\beta)mQ}dt. \tag{58}$$

Standard properties of Beta distributions, implies that an agent's posterior estimate is

$$\hat{p} = R(\alpha,\beta,m,f,g) \stackrel{\text{def}}{=} \lim_{Q\to\infty} \frac{J(\alpha,\beta,m,f,g,Q,1)}{J(\alpha,\beta,m,f,g,Q,0)}, \tag{59}$$

Moreover, taking the derivative of the term inside the integral of (58) with respect to t, and using the factor that for large Q, J converges to a scaled Dirac distribution, it follows that $\hat{p}$ satisfies:

$$\frac{\alpha}{\hat{p}} - \frac{1-\alpha}{1-\hat{p}} + f\left(\frac{\beta m}{g+f\hat{p}} - \frac{(1-\beta)m}{1-g-f\hat{p}}\right) = 0. \tag{60}$$

for large Q. Substituting in the equilibrium condition $\hat{p} = \bar{p}$, $\bar{p}$ as a function of r defined in (37), setting $\alpha = p$, $\beta = S_\tau(g + f\bar{p})$, $m = 1$, f and g as defined in (32), and solving for τ in (60) leads to the functional relation:

$$\begin{aligned}\tau(p,r) &= \Big(2p + 2p^2 - 32r + 20pr + 4p^2r + 36r^2 + 3pr^2 + \\ &+ 5p^2r^2 - 50r^3 - 5pr^3 + 3p^2r^3 + 25r^4 - 10pr^4 + p^2r^4\Big) \\ &/ \Big(-2p + 2p^2 + 20pr + 4p^2r + 48r^2 + 7pr^2 + \\ &+ 5p^2r^2 - 54r^3 - pr^3 + 3p^2r^3 + 15r^4 - 8pr^4 + p^2r^4\Big).\end{aligned} \tag{61}$$

This relation thus represents the level of visibility bias that is consistent with equilibrium, given p and r.

It is easy to verify that $\tau(p, 0) = 1$, and thus that the unbiased equilibrium with $r = 0$ is obtained in this case. Moreover, one verifies that τ is strictly increasing in r in a neighborhood of 0, regardless of p, and that τ approaches infinity for some $r < 1/2$, so equilibrium is defined for all parameter values p and τ, and r increases in τ, as does then $\bar{p}$. Finally, since $\bar{c}_0^I$ is decreasing in r, see (34), and $\bar{c}_0 = \bar{c}^e - \bar{c}_0^I$, it follows that $\bar{c}_0$ is increasing in r, and then also in τ, since r is increasing in τ.

∎

Further extensions

We first examine alternative utility functions. we then consider the model with arbitrary size of consumption bins.

The base model with other utility functions

The combination in the base model of the utility specification in (1), which leads to a linear consumption function in wealth, $\bar{c}_0 = \bar{p}\frac{W}{2}$, and the assumption that when $\frac{W}{2}$ is consumed at time 0 all bins contain consumption, makes the relationship between $\bar{p}$ and the expected number of consumption bin observations especially tractable, which allows for a strong characterization of equilibrium.

We now verify that qualitatively similar results as in Proposition 3 also hold under more common utility specifications. For example, consider the case in which agents have power utility, $U = \frac{c_0^{1-\gamma}}{1-\gamma} + \frac{c_1^{1-\gamma}}{1-\gamma}$, with risk aversion coefficient $\gamma \geq 1$ (where in the case $\gamma = 1$, log-utility is used). The consumption shock, ϵ is assumed to take on value $\frac{W}{2}$ with probability $1 - p$ (to avoid negatively infinite utility), and 0 with probability p. As before, the agent's estimated probability for a high outcome is $\hat{p}$

The agent's first order condition is in this case is

$$c_0^{-\gamma} = \hat{p}(W - c_0)^{-\gamma} + (1 - \hat{p})\left(\frac{W}{2} - c_0\right)^{-\gamma},$$

leading to the mapping $c_0 = G(\hat{p})\frac{W}{2}$. In the base model case with quadratic utility, $G(\hat{p}) = \hat{p}$. In the case of power utility, G is a nonlinear function for which a closed form solution is not available, bare a few special values of γ.[38] However, the following behavior of G is easy to show:

[38]For the special case when $\gamma = 1$, the closed form solution is $G(\bar{p}) = \frac{5}{4}\left(5 - \bar{p} - \sqrt{9 - 10\bar{p} + \bar{p}^2}\right)$.

Lemma A.3 *The function G satisfies $G(0) = \frac{1}{2}$, $G(1) = 1$, and is strictly increasing and convex. Its inverse is*

$$G^{-1}(c) = \frac{\left(1-\frac{c}{2}\right)^{\gamma}\left(\frac{c}{2}\right)^{-\gamma}\left(\left(\frac{c}{2}\right)^{\gamma}-\left(\frac{1}{2}-\left(\frac{c}{2}\right)\right)^{\gamma}\right)}{\left(1-\frac{c}{2}\right)^{\gamma}-\left(\frac{1}{2}-\frac{c}{2}\right)^{\gamma}}.$$

Proof: The form of G^{-1} follows immediately from the f.o.c. Differentiation of G^{-1} implies that G^{-1} is strictly increasing and concave on $c \in (1/2, 1)$. Moreover, $G^{-1}(1/2) = 0$, and $G^{-1}(1) = 1$. It follows that $G(0) = 1/2$, $G(1) = 1$, and from the inverse function theorem that G is invertible on $p \in (0, 1)$, being increasing and convex. ∎

If an agent observes a fraction x of consumption bins, the agent's posterior expected value of p is

$$\hat{p} = \frac{p + mG^{-1}(x)}{1+m}.$$

Owing to visibility bias, if other agents' consumptions are based on the posterior expected probability $\bar{p}$, then $x = S_\tau(G(\bar{p}))$. Finally, in equilibrium, $\hat{p} = \bar{p}$, leading to the following fixed point equilibrium condition:

$$\bar{p} = \frac{p + mG^{-1}(S_\tau(G(\bar{p})))}{1+m}. \tag{62}$$

The following proposition shows the existence of an equilibrium with over consumption in this setting:

Proposition A.1 *For $\tau > 1$, there exists an equilibrium probability estimate, $\bar{p} > p$, with associated consumption $G(\bar{p})\frac{W}{2} > G(p)\frac{W}{2}$.*

Proof: Note that $y = G(\bar{p}) \in \left(\frac{1}{2}, 1\right)$ satisfies

$$G^{-1}(y) = \frac{p + mG^{-1}(S_\tau(y)))}{1+m}. \tag{63}$$

To show the existence of a $y \in (G(p), 1)$ solving (63), we note that

$$p = G^{-1}(G(p)) < \frac{p + mG^{-1}(G(p))}{1+m} < \frac{p + mG^{-1}(S_\tau(G(p)))}{1+m},$$

and that

$$1 = G^{-1}(G(1)) > \frac{p + mG^{-1}(S_\tau(G(1)))}{1+m} = \frac{p+m}{1+m}.$$

By the intermediate value theorem, there is therefore a $y \in (G(p), 1)$ that solves (63), with associated equilibrium probability estimate $\bar{p} = G^{-1}(y) > G^{-1}(G(p)) = p$. ∎

The base model with arbitrary size of consumption bins

The assumption that an agent consumes in all bins when $c_0 = W/2$ makes the analysis tractable, since the calculus of Bayesian updates with Beta distributed priors and observations is straightforward. A generalization is to assume that when $c_0 = W/2$, a fraction $0 < f \leq 1$ of the bins are full. This allows us to analyze situations where there is heterogeneity in consumption behavior, for example, because of wealth uncertainty. Specifically, if a rich agent with probability estimate $\hat{p} = 1$ consumes in 100% of the consumption bins, then a poor agent with the same probability estimate will consume strictly less. The base model assumes $f = 1$, leading to the posterior estimate (6).

The following proposition covers the case when $f < 1$:

Proposition A.2 *The posterior expected probability of high consumption of an agent with prior p, who observes fraction z of bins being full, where each bin is full with probability pf, is*

$$\begin{aligned} \hat{p} &= R(p, z, m, f) && (64) \\ &= \frac{1}{2f(1+m)}\Big(1 + fp + fm + zm \\ &- \sqrt{(1 + fp + fm + zm)^2 - 4f(1+m)(p + zm)}\Big). \end{aligned}$$

Proof: Define

$$J(\alpha, \beta, m, f, Q, x) = \int_0^1 t^{\alpha Q - 1 + x}(1-t)^{(1-\alpha)Q-1}(ft)^{\beta m Q}(1-ft)^{(1-\beta)mQ}dt.$$

Using standard properties of Beta distributions, it follows that

$$R(\alpha, \beta, m, f) = \lim_{Q\to\infty} \frac{J(\alpha, \beta, m, f, Q, 1)}{J(\alpha, \beta, m, f, Q, 0)}, \tag{65}$$

and that an agent who updates according to Bayes rule will arrive at the posterior estimate $\hat{p} = R(p, z, m, f)$ when Q is very large. ∎

It is easy to verify that when $f = 1$, (64) reduces to the base model formula, $\hat{p} = \frac{p+mz}{1+m}$. Also, when $z = fp$, the formula reduces to $\hat{p} = p$, since the fraction of full bin observations is consistent with the prior in this case. Moreover, R is increasing in p and z, and is decreasing in f, since the lower f is, the lower the expected value of z is for a given prior p, which makes any given number z of observed full bins a more favorable indication about p.

Using similar arguments as before, an equilibrium probability estimate when visibility bias is present is then defined as a solution to the fixed point equation:

$$\bar{p} = R(p, S_\tau(\bar{p}f), m, f). \tag{66}$$

We now have

Proposition A.3 *There exists a unique equilibrium. In equilibrium there is overconsumption, and the equilibrium probability estimate is*

$$\bar{p} = B(1 + f(\tau - 1), p, m), \tag{67}$$

where the function B is defined in (19).

Proof: Substituting in the definition of R into the fixed point problem (66) yields a cubic equation in $\bar{p}$, two roots of which are outside of the unit interval $(0, 1)$. The remaining root has the prescribed form. ∎

The comparative statics from the base model therefore also hold in this variation. Moreover, increasing f has the same effect as increasing τ. Both lead to more overconsumption in equilibrium.

Corollary A.1 *The equilibrium probability estimate, $\bar{p}$ is increasing in the consumption fraction, $\partial \bar{p} / \partial f > 0$.*