Visibility Bias in the Transmission of Consumption Beliefs and Undersaving*

Bing Han† David Hirshleifer‡ Johan Walden§

October 22, 2018

Abstract

We study how bias in the social transmission process affects contagion of consumption beliefs and behavior. In the model, consumption is more salient than non-consumption. This visibility bias causes people to perceive that others are consuming heavily and to infer that others have favorable information about future wealth prospects. These inferences increase aggregate consumption. In contrast with other approaches, the visibility bias approach suggests that relatively simple disclosure policy interventions can ameliorate undersaving, especially when there are asymmetries of agent connectedness in the social network. In contrast with the Veblen wealth-signaling approach, information asymmetry about wealth reduces overconsumption. Our approach offers new implications about the effects of social connectedness, observation biases, and demographic structure; and offers a novel explanation for the dramatic drop in the savings rate in the US and several other countries in the last thirty years.

*We thank Pierluigi Balduzzi, Martin Cherkes, Marco Cipriani, Andreas Fuster, Giacomo de Giorgi, Mariassunta Giammati, Luigi Guiso, Robert Korajczyk, Theresa Kuchler, Chris Malloy, Timothy McQuade, Peter Reiss, Nick Roussanov, Andy Skrzypacz, Tano Santos, Siew Hong Teoh, Bernie Yeung, Youqing Zhou, participants at the American Economics Association meetings in Boston, 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 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.
§Haas School of Business, UC Berkeley.
1 Introduction

In acquiring attitudes about the world, people are heavily influenced by their cultural milieu, and by interactions with other individuals, especially when it is hard to come to a clear conclusion by introspection. Several authors have argued that people find it hard to decide how heavily to discount the future in making savings decisions (e.g., Akerlof and Shiller (2009)), owing either to lack of relevant information, or failure to process it effectively. It is hard to be sure what stream of satisfaction will actually result from a consumption/savings rule chosen today.\footnote{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.”} It is hard to forecast remaining lifespan or health in old age; most people do not process the relevant public but technical information contained in mortality tables and medical research. Finally, it is hard to predict risky future wealth realizations, and therefore, how much saving is needed today.

There is a great deal of evidence suggesting that people are indeed often ‘grasping at straws’ in their savings decisions. People make very basic mistakes, and rely on noisy cues, in deciding how much to save (Samuelson and Zeckhauser (1988), Shefrin and Thaler (1988), Madrian and Shea (2001), Beshears et al. (2008), Benhassine et al. (2015)). There is also considerable evidence that social interactions affect several dimensions of consumption, saving, and investment choices.\footnote{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).} Social interaction in general can have different implications from purely rational information transmission. These include contagion of unprofitable activities such as active trading in individual stocks; and contagion of behaviors whose advantages are publicly available information, such as participating in the stock market.\footnote{Individual investors are not, on average, good at trading individual stocks (Barber and Odean 2000).} Surprisingly, however, there has been little formal modeling of how bias in social learning processes affect consumption choices over time.

In our model, social learning about others’ consumption expenditures is biased by the fact that engaging in a consumption activity is more salient to others than not doing so. This leads to biases in inferring others’ beliefs about future wealth prospects. 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 acquires an enjoyable product as compared with not doing so. We call the greater availability and salience of potential
consumption events that do rather than do not occur visibility bias. We further assume
that people do not adequately adjust for the selection bias in favor of attending to the
consumption rather than nonconsumption events of others. This causes updating toward
the belief that others are consuming heavily.

These key premises of our model—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 consumption—are motivated by the psychology of
attention, salience, and social communication (see Section 2). With regard to communi-
cation, transmission of information in conversation is greater for positive information and for
more arousing information (Berger and Milkman 2012); we expect enjoyable consumption
in general to be more positive and arousing than the passive choice not to consume.

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 (or more generally, to salient events). However, failing to ad-
just appropriately for this selection bias in attention/observation is a clear mistake that
produces a directional bias in inferences.

Visibility bias makes consumption more available than non-consumption for later re-
trieval and cognitive processing. In consequence, people who naively neglect visibility bias
update toward thinking that others consume heavily and save lightly. So observers con-
clude that future consumption prospects are good and therefore that the low saving is
appropriate. Observers in our model therefore increase their own actual consumption.

In the model, the effects of visibility bias are self-reinforcing, as each individual becomes
an overconsuming model for others. So biased learning generates a positive feedback that
can result in severe undersaving in society as a whole, even when visibility bias is mild. In
market equilibrium, the reluctance of individuals to save implies a higher interest rate.

Visibility bias effects help explain a well-known puzzle in savings rates. 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 scholars and other observers about whether Americans will be able to sustain their
standards of living in retirement.4 A similar trend has occurred in many OECD countries,
with ratios of household debt to disposable income often reaching well over 100% (OECD
2014). Economists have long struggled to explain this drop (Parker 1999; Guidolin and
Jeunesse 2007).

come up with $400 to cover an emergency expense (of Governors of the Federal Reserve System 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).
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.” There are, however, a wide range of possible explanations, and we recognize that different scholars may have different perspectives about this conclusion.5

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. For example, the drop in costs of long-distance communication, 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. The rise of an increased diversity of cable television offerings (including channels devoted to shopping, travel, home remodeling, and other costly leisure pursuits, as well as dramas that less directly highlight consumption activities) further increased visibility. It is very common for people, in communicating by phone or other electronic networks, to report 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. Such biased observation is the driving force behind overconsumption in our model. Of course, there are many other drivers of savings rates, but visibility bias offers a possible explanation for the large and anomalous drop in saving.

There are also notable differences in savings rates across countries and ethnic groups which are not well-explained by traditional economic models (Bosworth 1993). An implication of the visibility bias approach is that relatively modest differences in beliefs or constraints can be amplified through social interactions. This can help explain the diversity of savings rates across groups. It suggests, for example, that cultural differences can have surprisingly strong effects.6 Our model also implies that degree of urbanization will be negatively related to savings rate, as urbanization is associated with a higher intensity of

5Existing explanations for the decline in savings rates include recent gains in household wealth coming from the stock market and real estate; upward revisions in individuals estimate of their permanent-income due to technological advances and higher productivity, growing income inequality (keeping-up-with-Jones effects), a stronger social safety net, increasing medical expenditure, an aging population, easier access to and improvements in the credit markets owing to financial innovations, greater present bias problems owing to rising availability of credit cards and mortgage credit, and trends in the way companies pay out to shareholders (disappearing dividends and more stock repurchases).

6Carroll, Rhee, and Rhee (1994) do not find an effect of culture on savings. They describe this as a tentative conclusion owing to data limitations. In contrast, using a similar methodology, Carroll, Rhee, and Rhee (1999) conclude that there are cultural effects.
social interaction and observation of the consumption of others. This is consistent with the evidence of Loayza, Schmidt-Hebbel, and Serven (2000), and is not directly predicted by non-social theories of consumption.

Overconsumption in our approach derives from mistaken beliefs about economic shocks and, in some cases, the behaviors of others. Such beliefs can potentially be corrected. So a distinctive empirical and policy implication of the visibility bias approach is that salient public disclosure about prospects for future consumption (which are, e.g., influenced by the risk of health and unemployment shocks), or about how much others actually consume, can help reduce overconsumption. Specifically, overconsumption in the model comes from overestimation of safety—people underestimate the probability of adverse wealth shocks. So vividly publicizing more accurate estimates about, e.g., risks of layoffs, or high health care bills, should reduce overconsumption.

However, in practice, announcements of probability estimates may be hard for people to process and convert into consumption plans. This suggests, as an alternative policy intervention, saliently disclosing information about how much others actually consume. Under some circumstances, such disclosures can also reduce overconsumption.

A literature in social psychology emphasizes that people tend to have biased perceptions about the attitudes and behaviors of others. For example, studies find that college students overestimate how much other students engage in and approve 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)). These studies argue that these misperceptions encourages such behavior, and that in some cases disclosure interventions can help remedy the problem.

In our model, agents observe upward-biased samples of the consumption of others. Agents neglect this selection bias, and update toward beliefs that favor higher consumption more than their priors would suggest. But surprisingly, in equilibrium actual consumption is so high that it confirms agents’ high expectations. So in contrast with the intuition from the abovementioned social psychology literature, the base version of our model demonstrates that social influence can induce more of a behavior without any overestimation of how much others engage in it. Since there is no overestimation of others’ consumption, disclosure of others’ average consumption does not correct agents beliefs.

However, some natural generalizations of the base model suggest conditions under which there will also, on average, be overestimation of others’ consumption. 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 setting with such “smart agents,” agents on average overestimate the consumption of others, and disclosure of actual average consumption reduces overconsumption. Furthermore, in reality
people interact in a social network in which some individuals are more heavily connected than others. When we allow for this, we show that there can on average be overestimation of others’ consumption, so that again disclosure of actual average consumption has a corrective effect.

Misestimation of the average actions of others also affects behavior in the model of Jackson (2018). In his framework, there is a strategic complementarity between agents’ actions, a possible example being recreational drug use. This means that a higher action level by one agent rationally leads to more of the action by others. In the version of the model in which agents do not account for the friendship paradox in their observations of others, agents overestimate how much others engage in the complementary action. This overestimation in equilibrium causes agents to engage more in the complementary behavior.

A key distinction between our model and Jackson’s is that there is no strategic complementarity in our setting—utility of consumption does not depend upon others’ consumption. Overconsumption in Jackson’s setting derives from overestimation of consumption complementarities. Overconsumption in our approach instead derives from agents’ naivete about visibility bias. So where Jackson’s approach is based on the interplay between naivete and payoff complementarities, our approach is purely belief-focused.

In the network extension of our model, the friendship paradox is present, further amplifying overconsumption. The fact that naivete about the friendship paradox can be an amplifier of certain behaviors is in the spirit of Jackson’s approach. However, even here, the mechanism works very differently. In Jackson (2018), highly-connected individuals (those with many social links) engage more heavily in the activity because their utility complementarities are stronger. In our setting, there are no such complementarities. Highly-connected agents are especially heavily influenced by visibility bias because they have more opportunities to observe their friends’ consumptions and update beliefs, causing greater overconsumption.

Although there is a large analytical literature on rational or biased social learning (e.g., see the survey of Golub and Sadler (2016)), and empirical literature on contagion of consumption or investment behaviors, there is surprisingly little research on whether social learning induces an overall directional bias toward over- or underconsumption. Empirical studies often focus on documenting the existence of contagion of specific consumption behaviors between individuals, rather than whether social interaction affects the average level of either specific behaviors or consumption as an aggregate. Similarly, we are not aware of theoretical research on this topic.

A plausible alternative theory of overconsumption and undersaving is that people are

---

7The friendship paradox in general agent networks refers to the fact that the average number of friends (connections) of agents’ friends is higher than the average number of friends agents have. The paradox arises generically, because well-connected agents are friends of more agents, and therefore disproportionately influential.
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 wealth dispersion matter, in contrast with pure individual-level biases such as present bias.

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 wealth signaling models, beliefs are rational, whereas the visibility bias approach is based upon biased inferences. The visibility approach has distinct empirical implications as well. For example, if all wealths were equal, Veblen effects would be eliminated, but the effects in the visibility bias approach still apply. So the visibility bias approach implies overconsumption even within peer groups with low wealth inequality. More generally, as discussed in Section 5.2, information asymmetry about other peoples wealths is the source of Veblen effects, whereas 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 investors deriving utility as a function of the consumptions of other investors (Abel 1990; Gali 1994; Campbell and Cochrane 1999). Unlike our model, the preference interaction approach does not in general result in equilibrium overconsumption (Dupor and Liu (1993), Beshears et al. (, forthcoming)). Concern for relative consumption as in ‘keeping up with the Joneses’ can induce a fear of falling behind which raises precautionary savings (Harbaugh 1996). Preference interactions potentially result in multiple equilibria with early or late consumption (Stracca and Al-Nowaihi 2005), whereas the visibility bias approach predicts a specific direction, overconsumption, along with other distinctive implications.8

As we have mentioned, a further distinctive empirical and policy implication of the visibility bias approach is that salient public disclosure helps correct people’s beliefs, reducing overconsumption. That a relatively simple policy intervention can potentially ameliorate the undersavings problem is specific to the visibility bias approach.

Finally, another approach that can lead to overconsumption is based on speculative disagreement (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

---

8Even with conventional preferences, externalities can also induce ‘keeping up with the Joneses’-like effects (DeMarzo, Kaniel, and Kremer (2004, 2008)), though models based on this approach do not focus on the issue of over- or under- consumption.
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 the speculative disagreement approach.

2 Psychology Background

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 consumption. We now briefly summarize evidence from the psychology of attention and salience that motivate and support our assumptions.

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 (2017)). 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 makes notable the phrase “The dog that did not bark” in the Sherlock Holmes story; it takes a genius to detect and recognize the importance of an absence. An example of the low salience of non-occurrences is neglect of opportunity costs, i.e., non-occurrences of benefits that would be received under alternative courses of action. Economics instructors are well aware that the opportunity cost concept is something that students struggle with. Neglect of absences is also reflected in the principle of WYSIATI, “What you see is all there is,” which Daniel Kahneman argues is one of the key features of System 1 thinking (Kahneman 2011).

There is evidence from both psychology, experimental economics, and field studies of selection neglect, a failure of observers to adjust appropriately for data selection biases (Nisbett and Ross 1980; Brenner, Koehler, and Tversky 1996). In general, neglect of selection bias is implied by the representativeness heuristic of Kahneman and Tversky (1972). Owing to limited cognitive resources, doing so requires time, 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. A model of how neglect of selection bias affects economic decisions is provided in Hirshleifer and Teoh (2003).

The combination of visibility bias and selection neglect in our model endogenizes the availability heuristic of Kahneman and Tversky (1973), so the tendency in the model to update toward thinking others are consuming heavily can be viewed as a consequence of

9People 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).
the availability heuristic. 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. For example, people overestimate the frequency of shark attacks because such attacks are vivid and heavily reported in the media relative to other causes of death such as car accidents. The availability heuristic is 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. In our model, this is disproportionately information about the consumption activities that were engaged in rather than not engaged in.

Experiments in the field confirm that engaging in a consumption activity is more salient to others than not engaging in the activity. As Frederick (2012) concludes, “purchasing and consumption are more conspicuous than forbearance and thrift.” For example, he comments that “Customers in the queue at Starbucks are more visible than those hidden away in their offices unwilling to spend $4 on coffee.” Our approach implies that this can result in overestimation of others’ consumption. Consistent with this, in Frederick’s experiment this salience results in overestimation by observers of how much other individuals value certain consumer products. Also consistent with visibility bias, people are influenced in car purchase decisions by observation of the purchases of others (Grinblatt, Keloharju, and Ikäheimo (2008), Shemesh and Zapatero (2016)), and such effects are stronger in areas where commuting patterns make the cars driven by others more visible (McShane, Bradlow, and Berger 2012).

Consumption activities of others may also be more cognitively available than non-consumption because someone who is consuming chooses to talk about it more than someone who is not consuming. Generally it is more interesting to hear about an action than inaction. 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. For various consumer products, positive word-of-mouth discussion of user experiences tends to predominate over negative discussion (see the review of East, Hammond, and Wright (2007)). A plausible reason is that users would like to persuade others of their expertise at product choice (Wojnicki and Godes 2008). This evidence suggests that people are more prone to sharing news about enjoyable consumption activities than the news that they stoically refrained from doing so.

3 The Basic Model

We first consider the effects of visibility bias in learning about the consumption of others on individual decisions when all investors have the same initial wealth and age. We start by focusing on individual optimization with interest rate given.
3.1 Optimal Consumption

Each individual maximizes a quadratic expected utility function with zero subjective rate of discount over two dates,

\[ U = c_0 - \left( \frac{\alpha}{2} \right) c_0^2 + E \left[ c_1 - \frac{\alpha}{2} c_1^2 \right]. \]  

At date 0, each individual chooses how much to consume and how much to borrow or lend at the riskfree interest rate \( r = 0 \). The budget constraint is

\[ c_1 = W - c_0 - \epsilon, \]

where, \( c_0 \) and \( c_1 \) are consumptions at dates 0 and 1, \( W \) is date 0 wealth, and \( \epsilon \) is a potential wealth shock at time 1. We assume that

\[ \epsilon = \begin{cases} 0 & \text{with probability } p \\ W & \text{with probability } 1 - p, \end{cases} \]

so with probability \( 0 < p < 1 \), the agent’s date 1 wealth is high, and with probability \( 1 - p \) it is low. We permit possible negative consumption, \( c_1 < 0 \). We assume that \( \alpha W < 1 \), to ensure that utility is increasing in consumption.

The negative wealth shock, which we view as being rare \( (1 - p \ll 1) \), can represent a systematic event, such as a major depression, or an underfunded pension system; or an idiosyncratic event that all agents are symmetrically exposed to, such as the possibility of a financially costly illness or disability, or a drop in demand for the agent’s human capital (e.g., job loss). The key is that agents draw inferences about the probability of such events (even if their occurrence is independent across agents) from their observations of the consumption of others.\(^{10}\)

The agent’s estimated probability that date 1 wealth will be high (\( \epsilon = 0 \)) is \( \hat{p} \). Based on this estimate, the agent chooses date 0 consumption to maximize expected utility, implying the first order condition

\[ 1 - \alpha c_0 = \hat{p}[1 - \alpha(W - c_0)] + (1 - \hat{p})[1 - \alpha(-c_0)]. \]

So optimal consumption is

\[ c_0 = \hat{p} \left( \frac{W}{2} \right), \]

\(^{10}\)An alternative modeling approach that would yield similar results would be to assume that agents are learning from others about the probability of dying young, which would also affect the benefits from saving. Yet another approach would assume that owing to visibility bias, people overestimate the subjective discount rates of others; and that owing to conformism, people update their own discount rates accordingly.
meaning that date 0 consumption is proportional to the estimated probability that future consumption will be high. It follows that if people were sure of a high outcome, they would consume half their total wealth; if \( \hat{p} < 1 \) they consume less than half. We do not model income (e.g., from human capital) in our model, but it is natural think of total wealth, \( W \), as being generated gradually, so that \( W/2 \) is generated in each of the two periods. It then follows that optimal saving (at time 0) is

\[
 s_0 = (1 - \hat{p}) \left( \frac{W}{2} \right). \tag{6}
\]

In this stylized model, the first and second periods are of equal length, leading to the half/half split of expected consumption. In a more realistic model the time periods could be viewed as unequal, which would result in other divisions of consumption.

Empirically, there is evidence that fear of adverse wealth shocks strongly affects consumption/savings decisions (Malmendier and Shen 2018). There is also evidence that people learn from the personal bankruptcies of neighbors about the risk of financial disaster, that learning of a neighbor’s bankruptcy causes people to cut back on their credit card expenditures, and that this results in a strong social multiplier effect (Agarwal, Qian, and Zou 2017).

### 3.2 Visibility Bias and Learning About Others’ Consumption

There are \( N \gg 1 \) identical agents, all facing identically distributed \( \epsilon \) risks. Total date 0 consumption of an agent is divided into \( K \) different activities which we call “bins” (\( K \) large), where each bin represents potential consumption of \( W/(2K) \). There are thus in total \( NK \) agent-consumption bins. An agent who chose to consume \( W/2 \) at date 0 (consistent with belief \( \hat{p} = 1 \), i.e., no risk of a negative shock) would then consume in every bin, whereas an agent who chooses to consume 0 (consistent with belief \( \hat{p} = 0 \), i.e., certainty of the adverse outcome) would not consume in any bin.

We refer to the agent’s prior as the agent’s perceived distribution of \( p \) based only on private signals, not social observation. The prior reflects private information, which makes it useful for agents to update based upon the observation of others. Specifically, each agent has a Beta-distributed prior for \( p \), which is based on observation of \( Q \) private signals about the level of \( p \). So \( p \sim Beta(Q_n, Q - Q_n) = Beta(Qq_n, Q(1 - q_n)) \), where \( Q \) is a common natural number for all agents, \( Q \in \mathbb{N} \), \( 0 \leq Q_n \leq Q \), and \( q_n \overset{\text{def}}{=} Q_n/Q \). Based on his

\[\text{In the cases } Q_n = 0 \text{ or } Q_n = Q, \text{ the prior is improper. The Beta distribution, which has support is } [0,1], \text{ is commonly used to describe the distribution of unknown probabilities of a Bernoulli random variable (in this case, the occurrence of the favorable wealth outcome). Intuitively, the agent’s prior is obtained by observing signals that are equivalent to observing } Q \text{ independent drawings of the Bernoulli variable, where } q_n \text{ is the fraction of successes (so that } Qq_n \text{ is the number of successes). So } Q \text{ is a proxy for}\]
signals, the agent’s prior estimate of the probability of a high outcome is then $q_n$. We call $q_n$ agent $n$’s prior type.

We may think of the prior type distribution as arising when each agent, starting with an improper $\text{Beta}(0, 0)$ “initial prior distribution” independently observes $Q$ consumption bins, each of which contains consumption with unbiased probability $p$, and conditions on these observations to update to a prior $\text{Beta}(Q_n, Q - Q_n)$ distribution, based on $Q_n$ bins with observed consumption. Viewing $Q_n$ as ex ante stochastic, it has a binomial distribution for agent $n$. It follows that since $N$ is very large, the fractions of agents of different prior types is deterministic and follows a binomial distribution across agents, as follows by the Glivenko-Cantelli theorem.

Specifically, each agent’s prior type is $\ell/Q$, where $\ell \sim \text{Binom}(Q; p)$. By the law of large numbers, it follows that the deterministic fraction $f_\ell$, of agents associated with prior type $\ell/Q$ is

$$f_\ell = \left(\frac{Q}{\ell}\right)(1 - p)^{Q - \ell}p^\ell, \quad \ell = 0, 1, \ldots, Q.$$

By a standard properties of binomial distributions over count variables (in this case, $\ell$), it follows that the average prior type is

$$\sum_{\ell=0}^{Q} \left(\frac{\ell}{Q}\right) f_\ell = p. \quad (7)$$

So on average, agents’ prior estimates are correct.

We refer to a bin as full if it contains consumption and empty otherwise. Each agent observes bins of the other agents’ $B = (N - 1)K$ bins, and believes the $M$ bins to be an unbiased sample.

The agent’s belief that he is observing a random sample may actually be incorrect. Crucially, we assume that observation is tilted toward those activities in which consumption did occur. This derives from what we call visibility bias, the tendency to notice and recall occurrences rather than non-occurrences. We view the event of engaging in a consumption activity as generally more salient to others than the event of not doing so.$^{12}$

One reason that consumption activities are highly visible is that many are social, such

---

11 The agent’s prior precision, and $q_n$ for the agent’s optimism.

12 The occurrence versus non-occurrence distinction that we focus upon is not the only source of differences in the salience of different consumption behaviors. Extreme outcomes also tend to be more 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- nor under-estimation of others’ consumption. Such effects (which we do not model) would basically be orthogonal to those we focus on. Our focus is on an attentional bias—neglect of nonoccurrences—that has a clearcut directional implication.
as eating at restaurants, wearing stylish clothing to work or parties, and traveling. Furthermore, physical shopping is itself a social activity. Both physical and electronic shopping and product evaluation are also engaging topics of conversation, either in person or online. In contrast, saving is often a private activity and investing are often undertaken privately through banks, brokers, or retirement account software. Many television dramas display glamorous consumption activities, travel, entertaining, and dining, and some media channels explicitly focus on shopping and other costly leisure activities. There are of course exceptions to these generalizations, such as investment clubs, but overall, consumption tends to be more observable and salient to others than is saving.

Specifically, if $B^C$ of the $B$ bins are full and $B^N$ are not, then we assume that the chance that an observed bin is full is

$$\frac{k^C B^C}{k^C B^C + k^N B^N} = \frac{B^C}{B^C + k^N \left(1 - \frac{B^C}{B}\right)} = \frac{x}{x + \frac{1-x}{\tau}} \overset{\text{def}}{=} S_\tau(x),$$

where $k^C$ is the probability that a bin is observed conditional upon it being full, $k^N$ is the probability that a bin is observed conditional upon it being empty, $\tau = \frac{k^C}{k^N} \geq 1$, and $x = \frac{B^C}{B}$ is the consumption fraction. The parameter $\tau$ measures the overrepresentation of full bins in the observer’s sample, i.e., visibility bias. When $\tau = 1$, the random observations match the actual distribution of consumption bins. When $\tau > 1$, there is overrepresentation of draws of consumption bins over non-consumption bins. The failure of the agent to adjust for this overrepresentation is a type of selection neglect—neglect of visibility bias. The number of full bins observed by agent $n$ is $Z_n$, $0 \leq Z_n \leq M$. We define $z_n = Z_n/M$ as the fraction of bins that are full, and $z = \frac{1}{N} \sum_n z_n$. Neglect of visibility bias, which tends to increase $z_n$, is the only deviation from rationality in the model. It is straightforward to show that the function $S_\tau(x)$ is strictly increasing in $\tau$ and $x \in (0, 1)$, is concave in $x$, and satisfies $S_\tau(0) = 0, S_\tau(1) = 1$, and $S_0(x) \equiv x$.

An agent uses Bayesian updating to estimate $p$ based on his bin observations, under
the belief that the fraction of full bins in the population is \( p \). Defining \( \xi = M/Q \), it follows that agent \( n \)'s posterior belief is

\[
\hat{p}_n = \frac{Qq_n + Mz_n}{Q + M} = \frac{q_n + \xi z_n}{1 + \xi}.
\]  

(9)

The variable \( \xi \) represents how much weight the agent puts on the new observations compared with his prior. This captures the intensity of an agent’s social interaction or social observation, and will be a source of some of the model’s distinctive empirical implications. An agent who observes others more updates more based on social observation.

Owing to visibility bias, agents tend to update heavily from their priors toward a belief that others have high consumption levels. The variables \( z_n \) and \( q_n \) for agent \( n \) are ex ante stochastic, but in the limit with many agents, the average estimate across agents \( \bar{p} = \frac{1}{N} \sum_n \hat{p}_n \), given the consumptions of all agents and the distribution of priors, is, by the law of large numbers,

\[
\bar{p} = \lim_{N \to \infty} \frac{1}{N} \sum_n q_n + \xi z_n = \frac{p + \xi E[z]}{1 + \xi}.
\]  

(10)

Each agent consumes in proportion to his probability estimate (as shown in (5)), and the average agent estimate is \( \bar{p} \), so this determines average consumption. It follows that the expected fraction of full bins observed by an agent under visibility bias is just \( S_\tau(\bar{p}) \), i.e.,

\[
E[z] = S_\tau(\bar{p}).
\]  

(11)

Specifically, each agent observes \( M \) bins, each with probability \( S_\tau(\bar{p}) \) of being full, leading to an expected observed consumption fraction of \( E[z] = E[z_n] = S_\tau(\bar{p}) \).

We define the mapping \( T \) from observation of consumption fraction \( S_\tau(x) \) to the posterior belief as

\[
T(x) \overset{\text{def}}{=} \frac{p + \xi S_\tau(x)}{1 + \xi}.
\]

An equilibrium is defined as a solution \( \bar{p} \) to (10,11), i.e., a fixed point \( \bar{p} = T(\bar{p}) \). It is easy to verify that the unique equilibrium \( \bar{p} \) when \( \tau > 1 \) is

\[
\bar{p} = \frac{(\tau - 1)(p + \xi) - 1 + \sqrt{V}}{2(1 + \xi)(\tau - 1)}, \quad \text{where}
\]

\[
V = [(\tau - 1)(p + \xi) - 1]^2 + 4p(1 + \xi)(\tau - 1).
\]  

(12,13)

\[\text{13}\]  

\[\text{15} \]  

All agents determine their consumption simultaneously. In a variation of the model, agents choose consumption sequentially based on (biased) observations of previous agents' consumption. The large sample equilibrium with sequential observations converges to the equilibrium with simultaneous observations that we study.
When $\tau = 1$, equilibrium is simply $\bar{p} = p$. There is no visibility bias, so agents are correct in their beliefs that the observed fraction of full bins in the population is $p$. When $\tau > 1$, agents mistakenly update as if $\tau = 1$. We write

$$\bar{p} = B(\tau, p, \xi)$$

for the function defined by (12,13) for $\tau > 1$, and by $B(1, p, \xi) = p$.

In equilibrium, different agents have different $\hat{p}_i$’s because of randomness in the number of Bernouilli successes built into each agent’s prior, $Q_n$ and the number of full bins observed socially, $Z_n$. But the aggregate estimate, $\bar{p}$, and corresponding per capita consumption, $\bar{c}_0$, are nonrandom by the law of large numbers. We call $\bar{p}$ the equilibrium probability estimate. It is proportional to aggregate (per-capita) consumption, $\bar{c}_0 = \bar{p} \left( \frac{W}{2} \right)$. The overconsumption factor, the ratio of consumption to optimal consumption (which is $p \left( \frac{W}{2} \right)$), is therefore $\bar{p}/p \geq 1$.

It is easy to verify that when $\tau > 1$,

$$\bar{p} > \frac{p + \xi S_r(p)}{1 + \xi}.$$  

In equilibrium an agent has higher consumption owing to visibility bias, thereby inducing higher consumption by other agents. This in turn encourages even higher consumption by the original agent. This feedback effect is reflected in the difference between the left-hand-side and right-hand-side of (15). As we shall see, the feedback effect can be powerful, especially when $\xi$ is high. The equilibrium has the following properties.

**Proposition 1** *In equilibrium:*

1. The equilibrium probability estimate, $\bar{p}$, and aggregate consumption are increasing in visibility bias, $\tau$, i.e., $\partial \bar{p}/\partial \tau > 0$, with $\bar{p} = p$ when $\tau = 1$.

2. As $\tau \to \infty$, $\bar{p} \to (p + \xi)/(1 + \xi)$, so that $0 < \bar{p}_\infty < 1$;

3. If $\tau > 1$, the average estimated probability of high consumption, $\bar{p}$, and aggregate consumption are increasing in the fraction of observations, $\xi$, i.e., $\partial \bar{p}/\partial \xi > 0$;

4. If $\tau > 1$, as the number of observations of others’ consumption bins, and thereby $\xi$, tends to infinity, $\bar{p} \to 1$, so that people ignore the risk of a bad outcome in determining their current consumption.

Intuitively, Part 1 says that owing to visibility bias in consumption observations, and neglect of sample selection bias (or equivalently, use of the availability heuristic) in assessing frequencies, people update heavily away from their priors toward a belief that others are
consuming heavily. In consequence, observers update too favorably about the information
others have about the probability of wealth non-disaster. This causes people to overcon-
sume, and the greater the visibility bias, the larger the effect.

Part 2 indicates that when visibility bias becomes maximally strong, beliefs become
maximally overoptimistic, but that agents’ prior beliefs have a moderating effect, so that
the equilibrium beliefs do not “spiral” upward toward $\bar{\rho} = 1$. Agents put some weight on
their priors, so even if 100% of observed bins are full, observers only update their beliefs
to $(p + \xi)/(1 + \xi) < 1$. The prior beliefs thus put an upper bound on the severity of
overconsumption.

Part 3 says that owing to visibility bias, greater observation of others as reflected in $\xi$
implies more optimistic beliefs and greater aggregate consumption. As biased observation
of others becomes very large relative to the prior precision, so that $\xi$ approaches infinity,
Part 4 says there may be drastic overconsumption (consuming as if they were sure there
were no risk of a bad outcome). New biased observations dominate prior information, so
that people become certain of a high outcome, even if visibility bias is small ($\tau$ close to one)
and the probability for a high outcome, $p$, is low. This is because when agents place heavy
weight on socially derived information, the feedback effect becomes very strong. Since $\bar{\rho} = p$
when $\tau = 1$, this also means that equilibrium consumption is very sensitive to changes in
$\tau$ for some parameter values (i.e., for large $\xi$). Together, Parts 2 and 3 suggest that the
feedback effect inherent in social transmission may be more important in generating severe
overconsumption than visibility bias itself.\textsuperscript{16} As a basic plausibility check, we also verify
that the average estimated probability of high consumption, $\bar{\rho}$, and aggregate consumption
are increasing in the true probability $p$ of no wealth disaster.

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 1 provide a possible
explanation.

Over the last several decades, improvements in electronic communications by such
means as phone (the drop in cost of long-distance telephone service), the rise of cell phones
and emails 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 $\tau$ and $\xi$, as in Parts 1 and 3 of Proposition 1. Greater obser-
vation and communication in general about the behavior of others is reflected by higher $\xi$
in the model. Greater opportunity to observe others intensifies the effects of visibility bias

\textsuperscript{16}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.
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 in favor of observing consumption over nonconsumption, i.e., visibility bias $\tau$. The activities that are noteworthy to report on very often involve expensive purchases, as with eating out or traveling. Indeed, 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 channel. 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 networking 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 would trigger a notification to friends; a non-posting about not engaging in a consumption event does not trigger a notification.

On special interest online discussion sites (e.g., focused on high tech or classical music), participants often post about associated product purchases. Such posting are more interesting, and therefore more likely to occur, than a posting to announce the news that the individual did not buy anything today.

So whereas electronic reports tend to select especially strongly for consumption activities, in-person unmediated observation of nearby individuals or close friends are likely to often include even nonconsumption activities. So the rise in modern communications results in an increase in visibility bias (i.e., larger $\tau$) and lead to higher overconsumption.\textsuperscript{17,18}

Past social research has also used other proxies for the intensity of social interaction and observation ($\xi$), such as population density (e.g., urban versus rural). 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.

The social influence parameter $\xi$ is identical across individuals. With diverse $\xi$’s, we

\textsuperscript{17}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.

\textsuperscript{18}Hirsh (2015) provides evidence that the drop in savings rates was accompanied by increasing population-level extraversion in many countries. Hirsh’s shifting extraversion explanation is compatible with our approach, since greater sociability causes greater observation of others’ consumption. However, even in the absence of shifts in population-level psychological traits, our model can explain the drop in the savings rates by improvements in communication technologies.
expect that individuals who engage in greater social observation will overconsume more than those with lower $\xi$. Such individuals update their beliefs more optimistically. A similar point holds for individuals who are more subject to visibility bias, i.e., greater $\tau$. It is evident that these predictions hold for the case in which $\xi$ or $\tau$ is identical for almost everyone.

**Proposition 2** Consider a society with common social observation parameter $\xi$ and visibility bias parameter $\tau$, with the exception of a deviant individual who has a social observation parameter value of $\xi'$, or alternatively a visibility bias parameter value $\tau'$. Then the expected consumption of the $\xi$-deviant is increasing with $\xi'$, and the expected consumption of the $\tau$-deviant is increasing with $\tau'$. A $\xi$-deviant on average consumes more than the others if and only if $\xi' > \xi$. A $\tau$-deviant on average consumes more than the others if and only if $\tau' > \tau$.

The result follows from (9), since

$$E[\hat{p}] = \frac{p + \xi' S_{\tau'}(\hat{p})}{1 + \xi'},$$

which is increasing in $\xi'$ and $\tau'$ (and a small deviant fraction does not alter the average probability estimate $\hat{p}$). Proposition 2 suggests that people who engage in greater social observation or are more subject to visibility bias will overconsume more. These implications are empirically testable. For example, survey data 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). Some studies have exploited information about actual social connections ((Heimer 2016), Bailey et al. ((2016, 2017)). For example, since neglect of visibility bias is an error, we expect $\tau$ to be higher for individuals and groups that are more subject to psychological bias, such as those with 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).

A possible objection to the conclusion of overconsumption is that houses serve as investment as well as consumption vehicles, and are highly visible to others. However, as discussed in footnote 13, the purchase of a house tends to be associated with an increase in the consumption of housing services, financed heavily by debt—i.e., an increase in current consumption at the expense of future consumption. Of course, buyers of expensive houses sometimes reduce non-housing consumption expenditures, but this need not imply an overall reduction in current consumption. Indeed, real estate equity is often accessed to finance non-housing consumption expenditures as well (Chen, Michaux, and Roussanov
A related objection is that in a multiperiod setting, the purchase of a house could be an indicator that an individual had saved heavily to accumulate enough for a substantial down-payment. We extend the model to allow for observation of the old by the young in Section 5.1 to allow for inferences about past saving, and find that in equilibrium, unambiguously, there is still overconsumption.

An interesting feature of the base model is that, despite naivety about visibility bias, agents end up with correct beliefs about others’ average consumption and beliefs. To see why, recall that all agents expect others to on average consume based on the correct value of $p$. In other words, agents do not foresee the overconsumption of others. It follows that agents would like to consume the same as the average consumption of their peers. Since all agents are ex ante identical, they have the same expected consumption, so on average they achieve this goal. Equivalently, their beliefs about $p$ are on average the same as others. This implies that on average they correctly assess the consumption of others.

This may seem counterintuitive, since agents are updating naively about the consumption of others based upon upward-biased samples. However, since agents think that others are not overconsuming, based upon priors, agents on average underestimate others’ equilibrium consumption. On average these effects exactly offset, so that people end up on average with correct assessments of others’ average consumption.\footnote{To put this another way, observers update from their priors toward on-average-high consumption observations of others, where these observations tend to be high even relative to others’ actual high consumption. But observers still place positive weight on their priors, so their updates are only partial. As is standard in Bayesian updating, they attribute the high level of the signal in part to randomness (or in the context of observing consumption bins, sampling error).}

### 3.3 Policy Interventions and Smart Agents

Overconsumption in our model derives from overestimation of safety from 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.

Such disclosure is effective in our model, but in practice it may be hard to make such disclosures salient and easy for people to interpret. Research on heuristics and biases consistently finds that people tend to put little weight on base rate probability information such as a numerical report about risk of layoffs (Kahneman and Tversky 1973; Borgida and Nisbett 1977). 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, and their effectiveness when there are ‘smart agents.’ In Subsubsection 3.3.1 we consider disclosures of the average consumption or saving rate of peers, where there is no visibility bias in the disclosure. By this we mean that the disclosure does not do anything special to highlight consumption versus nonconsumption activities, nor the actions of those who are consuming heavily versus those who are saving heavily. In Subsubsection 3.3.2, we consider such visibility-biased disclosures.

3.3.1 Public Disclosure of Average Consumption of Peers

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. Since people in our approach do not take into account that others are biased, under plausible variations of the base model assumptions, as we shall see, people may end up with biased perceptions about what others believe and how much they consume. If so, saliently publicizing accurate information about others can help alleviate overconsumption. Specifically, if people overestimate how optimistic others are, and how much others consume, then accurate information about others would correct that mistake, reducing overconsumption.

Empirically, in tests covering a very wide range of activities, the intervention of providing accurate information about peer beliefs or behavior tends to cause behavior to conform more closely to the disseminated peer norm. Studies in which peer information promotes conformity toward the disclosed norm include 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), and Chen et al. (2010).20

Social norms marketing, or dissemination of information about one’s peers, can be an effective policy tool for correcting inaccurate beliefs about peers, and to makes peer actions more salient.

In an example of the effects of peer disclosure, informing college freshmen of survey results about the attitudes of other college students toward heavy drinking is associated

---

20However, 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 (in the context of our model, e.g., not having identically distributed wealth shocks). So our theory does not make a prediction about the outcome of this experiment.
with lower levels of self-reported drinking 4-6 months later, as well as lower acceptance of pro-drinking norms (Schroeder and Prentice 1998). The explanation offered by the authors was that students were conforming to a better and more realistic norm. In a recent test, Bursztyn, González, and Yanagizawa-Drott (2018) report that a very large majority of young married men in Saudi Arabia support female labor force participation outside of home, but substantially underestimate how much similar men support this. In an incentivized experiment, randomly correcting beliefs about other men increases men’s willingness to let their wives join the labor force, and actual labor force participation by their wives. In the context of our model, these findings suggest that salient reporting of actual consumption of peers may reduce consumption.

We have seen that in the base model as developed so far, agents on average have correct assessments of the average consumption of others. It follows that disclosure of the actual average consumption of others will not change the average level of overconsumption. Specifically, an agent with prior $q_n$ and bin observations $z_n$, who observes a publicized signal of $\bar{p}$ (or equivalently of aggregate consumption), arrives at the posterior

$$\hat{p}_n = \frac{q_n + \xi z_n + \alpha \bar{p}}{1 + \xi + \alpha}.$$  

Here, the parameter $\alpha \geq 0$ determines how much weight the agent puts on the public signal. The equilibrium condition then becomes

$$\bar{p} = \frac{p + \xi S_r(\bar{p}) + \alpha \bar{p}}{1 + \xi + \alpha}, \quad (16)$$

which, as is easily verified, leads to the same equilibrium as in the base model (which corresponds to $\alpha = 0$).

However, in a slight generalization of the base model, there is systematic overestimation of the consumption of others. Suppose that there is a group of more knowledgeable or sophisticated agents (“smart agents”). Specifically, we now allow for a second group of agents consisting of fraction $\phi$ of the population who end up with unbiased expectations, i.e., their average posterior belief is $p$. This would occur, for example, if these are smart agents who rationally adjust for visibility bias in their observations. Alternatively, even if

---

21 In a large population, a perfectly accurate disclosure of the average behavior would reflects an extremely large sample. This would be so informative that agents would put arbitrarily heavy weight upon the public signal. However, in reality empirical proxies for aggregate consumption are noisy, so that a Bayesian (or quasi-Bayesian) would put only a finite weight on it. Noise would also be present if each agent were not completely sure whether the population the policy-maker discloses about matches the agent’s relevant set of peers (people with similar preferences and endowments, for example). For simplicity, instead of modeling noise in detail, we assume that each agent places only a finite weight $\alpha$ on the disclosed average consumption of others.
these agents are subject to visibility bias, this outcome will occur if they have strong prior information about \( p \) (i.e., the number of signals \( Q \) reflected in their priors is arbitrarily large). For example, these agents might have studied the statistics on the frequency of expensive illness or of job loss. Then their beliefs will tend to be close to \( p \), and they won’t update much when they learn the population average consumption. The remaining \( 1 - \phi \) agents, just as before, neglect visibility bias.

Now the average population belief \( \bar{p} \), and therefore also aggregate consumption, reflect a balance of beliefs between the smart agents and the other agents (whom we will refer to as ‘biased’). Let \( \bar{p}^V \) be the average belief of the biased agents (\( V \) for visibility bias). Intuitively, in the absence of public disclosure, the average belief is dragged down by the smart agents (beliefs near \( p \)), and dragged up by the optimism of the biased agents (\( \bar{p}^V > p \)). This suggests implies that the beliefs of the biased agents are above average (\( \bar{p}^V > \bar{p} \)). So a salient disclosure that indicates that the average belief is \( \bar{p} \) pulls down the beliefs of the biased agents substantially, and only modestly pulls up the beliefs of the smart agents. So overall \( \bar{p} \) declines, reducing overconsumption.

To formalize these intuitions minimally, consider the case in which the smart agents (whom we have described as either rational or as well-informed) know the true \( p \). Then the biased agents observe \( S_r(\phi p + (1 - \phi)\bar{p}^V) \). In updating, just as above, they also put weight \( \alpha \) on the public disclosure of \( \bar{p} \). So equilibrium with the public signal can be defined as the modification of (16),

\[
\bar{p}^V = \frac{p + \xi S_r(\bar{p}) + \alpha \bar{p}}{1 + \xi + \alpha}, \quad \bar{p} = \phi p + (1 - \phi)\bar{p}^V.
\]

This requires that the average belief of the biased agents is the average of the updated beliefs that such agents form based on their naivete about the biased samples they observe, and based on the public signal. It follows that biased agents believe that average consumption is higher than it is, \( \bar{p}^V > \bar{p} = \phi p + (1 - \phi)\bar{p}^V \). Comparing the case case \( \alpha > 0 \) with the no-disclosure case of \( \alpha = 0 \), we see that in the smart agents model, publicizing a (potentially noisy) signal about the average belief \( \bar{p} \) or average consumption of others in equilibrium decreases aggregate consumption.

The above discussion implies:

**Proposition 3** Consider the modification of the base model from Section 3.2 in which there is a fraction \( \phi \) of ‘smart’ agents who know the true probability \( p \) of no disasters, and where average consumption is publicized. Then

- Agents on average overestimate the average consumption of others.
- Average consumption is decreasing in the weight, \( \alpha \), that agents with visibility bias assign to the public signal.
We are not aware of evidence about whether people tend to overestimate the aggregate consumption of others (as in this smart agents model), or not (as in the base model equilibrium). Which of these is the case has important policy implications.

The conclusions of Proposition 3 would also hold even if ‘smart’ agents were not as smart, as long as their average belief, in the absence of disclosure, is below the population average, and they are sufficiently resistant to updating upward in response to disclosure of the population average. In the proposition, these two features are achieved by having all smart agents know the true \( p < \bar{p} \). This certain knowledge makes them resistant to updating upward after disclosure.

Another possibility is that smart agents are investors who are overconfident in the sense that they mistakenly have more faith in the accuracy of their own signals than in the signals of others.\(^{22}\) As such, they are little influenced by their visibility-upward-biased personal observations of others, causing their average beliefs to be below the population average. Furthermore, when the public signal about average consumption is disclosed, they place little weight on this information as well, so they do not revise their beliefs upward very much. Another pathway to the result is if ‘smart’ agents are smart only in the sense that they understand that other agents neglect visibility bias. If so, then smart agents will recognize that others overconsume, and hence smart agents will not update upward as strongly as other agents do based upon their upward-biased personal observations of others. So again, the beliefs of smart agents will be below the population mean. Furthermore, the disclosure of the population mean has little effect on smart agents, since they understand both before and after the disclosure that others are unduly influenced by visibility bias.

When we consider our extension to a setting with overlapping generations, we will see why a different kind of disclosure—of the consumption of only a subset of the population—can also help address the problem of overconsumption.

A different possible approach to correcting overconsumption, which we explore in the next section, is to fight visibility bias with visibility bias. If there is a way to make saving behavior more salient, observers who neglect visibility bias will tend to update more toward a belief that others save heavily and that saving is desirable. This approach can be effective even when there are no smart agents in the model, and no misperceptions about others’ average consumption.

\(^{22}\)We consider here overconfidence about priors that are, on average unbiased. I.e., we are not referring here to systematic overoptimism. Overconfidence in the sense of overestimating the accuracy of one’s own beliefs is extremely well-documented in the psychology and economics literatures (see, e.g., the surveys of Daniel and Hirshleifer (2015), Moore, Tenney, and Haran (2015)). It is also well-documented that people are heterogeneous in their degrees of overconfidence.
3.3.2 Visibility-Biased Disclosures

We now examine disclosures that are subject to visibility bias, in the sense that the disclosure focuses on consumption versus nonconsumption behaviors. We first examine analytically the effect of accurate public disclosures that make the saving behavior of others more salient. We verify that in equilibrium such visibility-biased disclosures encourage saving. We then discuss in practical terms how to make a disclosure that highlights saving behavior, and how effective such disclosures are likely to be. Since the effects we describe here do not require smart agents, we return to the base assumption that all agents are ex ante identical.

As we have just seen, when a public disclosure of a signal about \( \bar{p} \) (or equivalently, about consumption) is added to the base model, see equation (16), equilibrium consumption is unaffected and that when there are smart (or stubborn unbiased) agents, equilibrium overconsumption declines. This analysis is based on a plain-vanilla disclosure that does not draw disproportionate attention to either consumption or nonconsumption (i.e., saving) behaviors.

Suppose instead that the signal calls attention more to nonconsumption than to consumption behavior.\(^{23}\) We will refer to this as a signal about “saving,” but clearly in substantive terms, a disclosure about saving is equivalent to a disclosure about consumption, since one can infer one from the other. Consumption is proportional to \( \bar{p} \), and saving to \( 1 - \bar{p} \) (see equations (5) and (6)). Observing a signal about average saving is thus equivalent to observing a signal about \( 1 - \bar{p} \). So we are now not just referring to any accurate disclosure that is informative about the amount of saving (or equivalently, consumption); we are referring to a disclosure that makes salient the saving rather than consumption activity of others.

What does it mean to make an accurate disclosure which is biased toward visibility of saving rather than consumption? People could be given stickers that they are free to post on their cars or in personal spaces, saying “Proud Retirement Saver.” The policymaker could have an advertising campaign explaining that these signs or stickers are given to anyone who is saving more than some prespecified absolute amount or fraction of income. There is visibility bias, since the presence of the sign or sticker is more noticeable than the absence of one.

In the model, if there were no visibility bias to the saving disclosure, in (16) above, agents would infer that a signal about saving (basically \( 1 - \bar{p} \)) is equivalent to a signal about consumption (basically \( \bar{p} = 1 - (1 - \bar{p}) \)), and of course the conclusion would remain unchanged. However, if there is visibility bias in the disclosure toward the occurrence of saving behavior, and observers neglect this visibility bias, then observed saving is overes-

\(^{23}\)Importantly, we are only examining accurate disclosures. They direct attention to saving behavior, but are not misrepresenting the amount of it that is occurring.
timated, so that the visibility bias function $S_\tau$ is applied to the amount saved. It follows that equilibrium beliefs in this case satisfy

$$\bar{p} = \frac{p + \xi S_\tau(\bar{p}) + \alpha(1 - S_\tau(1 - \bar{p}))}{1 + \xi + \alpha}. \quad (18)$$

The $1 - S_\tau(1 - \bar{p})$-term in the expression reflects the agent’s neglect of visibility bias about saving. The agent believes that he is observing a signal about $1 - \bar{p}$ (proportional to saving), when he is actually observing a signal about the visibility-biased quantity $S_\tau(1 - \bar{p})$. This leads to the inferred belief (proportional to consumption) of $1 - S_\tau(1 - \bar{p})$, rather than $\bar{p} = 1 - (1 - \bar{p})$.

Equation (18) has a unique closed form solution, and it is not hard to derive the following:

**Proposition 4** Equilibrium consumption when there is visibility bias toward higher saving, as given in (18), is lower than when there is no visibility bias toward higher saving, as given in (14).

Publicizing the visibility-biased signal about savings thus unambiguously leads to lower consumption. In other words, the disclosure fights visibility bias with visibility bias. Neglect of visibility bias about the disclosure encourages saving, by partly offsetting the effects of visibility bias in agents’ direct observations of others’ consumption activities.

For several reasons, interventions by policymakers to make saving behavior more salient, as reflected in the $\alpha$ term in the numerator of equation (18), is unlikely to fully offset the spontaneous visibility bias toward observing the consumption activities of others, as reflected in the $\xi$ term. People are heavily exposed to the consumption of others many times each day as people interact others in-person or electronically. In contrast, policy campaigns are likely to be episodic and to generate a relatively limited number of observations. Nevertheless, if the campaign succeeds in making observations about others’ saving behaviors sufficiently salient, there can be a beneficial effect.

A drawback of the sticker intervention that we have described is that people have little incentive to post the stickers. Indeed, some might fear that posting such a sticker would look like bragging. However, a more nuanced scheme could potentially be designed to minimize these problems.\(^{24}\)

\(^{24}\)For example, the scheme could be localized, with members of a neighborhood receiving stickers saying “Proud Retirement Saver.” Signs would be set up in the neighborhood as locations for people to post their stickers anonymously, but people would also be encouraged to post them on their own cars, or at their homes or workplaces. The policymaker counts the number of stickers posted on the public sign posts; people self-report about posting in personal locations. Savers as a group are given monetary bonus when there are more posted stickers in the area. Again, the presence of the stickers leads to a clear visibility bias toward saving over consumption.
4 Observation of Others in a Social Network

We can describe an agent’s linkages in a social network in terms of whose consumption an agent can potentially observe. An agent’s location in the network can affect perceptions of the consumption of others. So it is interesting to study how network location affects consumption.

We now extend the model to allow for an arbitrary social network, to derive empirical implications for how individual centrality and overall network connectedness affect beliefs and consumption.

Agents are connected in an undirected social network represented by the graph \( G = (\mathcal{N}, \mathcal{E}) \), where \( \mathcal{N} \) is the set of investors and \( \mathcal{E} \) is the set of edges connecting them. The set of agents \( \mathcal{N} = \{1, \ldots, N\} \), and \( (m, n) \in \mathcal{E} \subset \mathcal{N} \times \mathcal{N} \) if investors \( m \) and \( n \) are connected through a direct social tie. By convention, the network is undirected, i.e., \( (m, n) \in \mathcal{E} \Leftrightarrow (n, m) \in \mathcal{E} \), and investors are not connected to themselves \( ((n, n) \notin \mathcal{E}) \). The set of agents that \( n \) is socially linked to is \( \mathcal{D}_n = \{m : (n, m) \in \mathcal{E} \subset \mathcal{N} \setminus \{n\} \} \), and \( n \)'s connectedness is \( d_n = |\mathcal{D}_n| \). The maximal degree in the network is \( D = \max_n d_n \).

Associated with the network is the symmetric adjacency matrix \( \mathbf{E} \in \mathbb{R}^{N \times N} \), with \( \mathbf{E}_{mn} = 1 \) if \( (m, n) \in \mathcal{E} \), and \( \mathbf{E}_{mn} = 0 \) otherwise. We focus on a connected network (meaning that there is a path between any two agents). Each agent therefore has at least one neighbor.

Agent \( n \), with prior type \( q_n \), randomly observes \( d_n M \) consumption bins of his neighbors’ \( d_n K \) bins. Here, we assume that \( K \) is sufficiently large that all agents treat these observations as effectively independent, i.e., as if the agent were sampling with replacement. An agent with more neighbors will thus have more observations, and therefore update his consumption behavior more aggressively than an agent with few neighbors. This is captured by the variation of \( d_n \), as contrasted with the base model in which all agents have the same number of observations, \( M \). Given their observations of their neighbors, each agent forms posterior beliefs \( \hat{p}_n \), that govern their own consumption. Specifically, \( z_n \) is the fraction of agent \( n \)'s observations of \( n \)'s neighbors’s bins that are full. It follows that when there is visibility bias, \( \tau > 1 \),

\[
E[z_n] = S_\tau \left( \frac{1}{d_n} \sum_{m \in \mathcal{D}_n} \hat{p}_m \right).
\]

\(^{25}\)To ensure that there is a large enough number of agents so that the law of large numbers can be used, as in the previous section, we make the technical assumption that there are actually a large number of agents representing each node position in the network. Each agent randomly observes the consumption bins of agents in its neighboring node positions. The approach is similar to the replica network approach in Walden (2018). Each node in the network thus represents a whole equivalence class of identical agents, and there is a sufficient large number of agents in the economy so that expectations rather than realizations can be used in the subsequent equilibrium fixed point definition, as in the base model of Section 3.2.
Definition 1 A network consumption equilibrium is a vector, \( \vec{p} = (\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N)' \in [0, 1]^N \), such that

\[
\vec{p}_n = \frac{p + d_n \xi S_\tau \left( \frac{1}{d_n} \sum_{m \in D_n} \vec{p}_m \right)}{1 + d_n \xi}, \quad n = 1, \ldots, N.
\]

This is the natural generalization of the equilibrium concept used in the preceding section.

The per capita equilibrium probability estimate is

\[
\bar{p} = \frac{1}{N} \sum_n \vec{p}_n,
\]

and the per capita consumption is \( \bar{c}_0 = \bar{p} (W/2) \). The network economy is characterized by the tuple \( T = (E, \xi, \tau, p) \), where \( E \) is the adjacency matrix of the connected network, \( \xi > 0, \tau > 1, \) and \( 0 < p < 1 \). Owing to neglect of visibility bias, each agent behaves as if the economy is actually \( T' = (E, \xi, 1, p) \).

The following proposition characterizes the equilibrium:

**Proposition 5** Consider an economy represented by \( T = (E, \xi, \tau, p) \).

1. There exists a network equilibrium vector, \( \vec{p} \in [p, 1]^N \), with correct consumption, i.e., \( \vec{p} = (p, p, \ldots, p)' \) if and only if \( \tau = 1 \).

2. The equilibrium vector is unique if

\[
\left( 1 + \frac{1}{\xi D} \right) [(\tau - 1)p + 1]^2 > \tau. \quad (21)
\]

Alternative sufficient conditions for uniqueness are that:

(i) The prior probability for high consumption is sufficiently high, \( p > \frac{1}{2} \), or

(ii) Visibility bias is sufficiently low, i.e., \( \tau \) is sufficiently close to one, so that \( \tau < 1 + \frac{1}{\xi D} \), or

(iii) Visibility bias is sufficiently high, i.e., \( \tau \) is sufficiently large, so that \( 1 < p(\tau - 1) \).

3. As \( \tau \to \infty \), the equilibrium vector converges to

\[
\vec{p} = \left( \frac{p + d_1 \xi}{1 + d_1 \xi}, \frac{p + d_2 \xi}{1 + d_2 \xi}, \ldots, \frac{p + d_N \xi}{1 + d_N \xi} \right)'.
\]

From here on, we focus on the case when \( p > 1/2 \), justified by our assumption that the negative wealth shock (which occurs with probability \( 1 - p \)) is quite rare. This implies that the equilibrium vector is unique.
For reasons of tractability, network models often focus on symmetric networks. In this context we define a network as symmetric if all agents having the same connectivity, \( d \). For symmetric networks, we also have:

**Proposition 6** When the social network is symmetric, equilibrium satisfies \( \bar{p}_n = B(\tau, p, d\xi) \), \( n = 1, \ldots, N \), so that all agents have the same beliefs and consumption.

Specifically, in a symmetric network, all agents consume \( \left( \frac{W}{2} \right) B(\tau, p, d\xi) \). It follows that all the results in Proposition 1 generalize to symmetric networks. Moreover, the following results hold with respect to connectivity, \( d \):

**Corollary 1** When the social network is symmetric:

- Equilibrium consumption \( \bar{c}_0 \) is increasing in connectivity, \( d \).
- As connectivity, \( d \to \infty \), equilibrium consumption approaches \( W/2 \), corresponding to \( \bar{p}_n = 1 \) for all \( n \).

So, overconsumption is more pronounced in more well-connected societies.

### 4.1 Individual consumption and centrality

The network equilibrium relation (20) suggests that an agent’s equilibrium consumption increases in his connectedness, \( d_n \), because the more connections he has, the more weight he puts on his observations and the less weight on his (lower) prior. This is the only mechanism through which an agent’s number of connections is important in our model. This distinguishes our model from that of Jackson (2018), which assumes a direct payoff complementarity wherein the incremental payoff from engaging in the behavior is increasing in the product of the number connections an agent has and the average behavior of those connections.

In our setting, an agent’s consumption is also higher the higher the consumption of the agents he is connected to. The consumption of these neighbors, in turn, tends to be increasing in their connectedness. So an agent’s consumption depends upon a potentially unlimited iteration of dependencies, where each stage tends to be increasing with the relevant agents’ connectedness.

Measures of *centrality* from network theory are sometimes designed to take into account such iterated dependencies. This suggests that agents that are more central (well-connected) will overconsume more. This can only be evaluated in a network where agents

---

26 Most of the many ways of defining network symmetry impose stronger conditions than we do (e.g., that the number of nodes at distance \( t \) is the same for any two nodes and any \( t \geq 1 \)).
differ in their connectivity. To examine such effects, we therefore now consider asymmetric
networks.

We study a class of networks in which there is a core of highly connected agents sur-
rounded by peripheral, less connected, agents who are mainly connected to the core. In
a social context, we may think of the networks core as highly social agents, with lots of
connections among themselves and to others. Such networks are thus asymmetric.\footnote{Core-periphery networks arise in many different real world contexts, e.g., in over-the-counter dealer networks.}

For tractability, we study core-periphery networks in which all core agents have the same
number of connections to other core agents, namely $d^C > 0$, and also the same number,
$d^P > 0$ to peripheral agents. Each peripheral agent is connected to only one core agent.
An example of a network with $d^C = 3$, $d^P = 3$ is shown in Figure 1. Note that $d^P$ also

denotes the number of peripheral agents per core agent in the economy.

It follows from Definition 1 that equilibrium probability estimates of the core and pe-

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{core-periphery-network.png}
\caption{Core-periphery network, with 4 (black) core agents in center and 12 (red) peripheral agents. Each core agent has $d^C = 3$ connections to other core agents, and $d^P = 3$ connections to peripheral agents.}
\end{figure}
ipheral agents, $\bar{p}^C$ and $\bar{p}^P$, satisfy

$$\bar{p}^C = \frac{p + (d^C + d^P)\xi S_T \left( \left( \frac{d^C}{d^C + d^P} \right) \bar{p}^C + \left( \frac{d^P}{d^C + d^P} \right) \bar{p}^P \right) }{1 + (d^C + d^P)\xi},$$  \hfill (22)

$$\bar{p}^P = \frac{p + \xi S_T (\bar{p}^C)}{1 + \xi}. \hfill (23)$$

Also, the per capita consumption in the economy is $\bar{c}_0 = \bar{p} \left( \frac{W}{\tau} \right)$, where

$$\bar{p} = \left( \frac{1}{d^P + 1} \bar{p}^C \right) + \left( \frac{d^P}{d^P + 1} \right) \bar{p}^P$$

is a weighted average of the individual agents’ probability estimates. Under our assumption that $p > 1/2$, by Proposition 5, the equilibrium is unique. The following result shows how the structure of the core-periphery network determines equilibrium overconsumption.

**Proposition 7** Core agents consume more than peripheral agents, $\bar{p}^C > \bar{p}^P$, and aggregate consumption is increasing in the connectivity of core agents, $d^C$.

Proposition 7 shows how the composition of the core-periphery network influences the amount of aggregate overconsumption in the economy. When the core agents are more heavily connected, there is greater overconsumption. So if social trends, such as the rise of Facebook or Twitter, result in core agents becoming more heavily connected, we expect overconsumption to increase.

### 4.2 Social Networks and Policy Interventions

We have seen in Subsection 3.3 that when there are smart agents, on average agents overestimate the average consumption of others. It followed that salient public disclosure of the consumption of others helps reduce overconsumption.

We will now see that even without smart agents, when we make the realistic assumption that the social network is asymmetric, again there are misperceptions of the beliefs and behavior of others. In consequence, the disclosure of public information about what others think and do affects overconsumption.

We will illustrate this in a simple example in which agents end up overestimating the consumption of others. Consider a variation of the network model, in which all agents receive an additional unbiased signal about aggregate consumption, by observing $\bar{p}$ (or an unbiased noisy version thereof), which they incorporate into their Bayesian posteriors. This leads to the following definition of network equilibrium:
Definition 2  A network consumption equilibrium with a public signal about aggregate consumption is a vector, $\bar{p} = (\bar{p}_1, \bar{p}_2, \ldots, \bar{p}_N)' \in [0, 1]^N$, such that

$$\bar{p}_n = p + d_n \xi S_{\tau} \left( \frac{1}{d_n} \sum_{m \in D_n} \bar{p}_m \right) + d_n \alpha \bar{p} + \frac{d_n \xi d_n \alpha}{1 + d_n \xi + d_n \alpha}, \quad n = 1, \ldots, N,$$

(24)

where $\bar{p} = \frac{1}{N} \sum_n \bar{p}_n$.

In this variation, agents’ posteriors are based on three components: their priors, their observations, and the public signal. As before, the parameter $\alpha \geq 0$ determines how much weight the agents put on the public signal. When $\alpha = 0$, the extension reduces to the original network model.

We saw in the base model, without smart agents, agents on average correctly assess the average consumption of others. It follows that the presence of a public signal about aggregate consumption does not correct overconsumption. A similar point applies in our network setting if the network is symmetric.

Proposition 8  In any symmetric network, aggregate consumption is unaffected by the weight agents assign to the public signal, $d\bar{p}/d\alpha = 0$.

This comes from the fact that agents consume based on posterior beliefs, which on average are equal to $\bar{p}$. It follows that on average the public signal just reconfirms these average posterior beliefs, and therefore has no effect on aggregate consumption. The intuition for this result is essentially the same as the intuition for why agents correctly assess the consumption of others in the base model.

Proposition 8 provides conditions under which publicizing aggregate consumption data does not help address overconsumption. However, this relies on the crucial assumption that networks are symmetric. A key property of symmetric networks in our model is that the average posterior belief is consistent with actual aggregate consumption. This is in general not the case in asymmetric networks. As previously analyzed, in our model central agents in asymmetric networks consume more, and also disproportionally influence the consumption of other agents. Therefore, the consumption of the average observed agent tends to be higher than the consumption of the average agent. This point is a reflection of the majority illusion in social networks (Lerman, Yan, and Wu 2016), wherein observers disproportionately see the characteristics of better-connected agents. (This effect is basically a generalization of the friendship paradox, which was discussed in the introduction.) The higher average consumption of observed agents in an asymmetric network makes the public signal informative. The public signal is below that of the average observed agent, so the public signal acts as a corrective to overconsumption.
We verify this effect in the core-periphery network in the previously studied case, where it is easy to verify that \( \bar{p} \), and thereby also average consumption, is decreasing in the weight assigned to the public signal.

**Result 1** In a core-periphery network with \( d^C = 3 \), \( d^P = 3 \), \( \xi = 1 \), \( \tau = 2 \), and \( p = 3/5 \), it is easy to verify that \( \bar{p} \), and thereby also average consumption, is decreasing in the weight assigned to the public signal, \( \alpha \).

To sum up, just as in the ‘smart agents’ setting of Subsection 3.3, in an asymmetric networks setting, a public signal helps expose mistaken beliefs that agents have about the consumption of others. This makes public perceptions more accurate and reduces overconsumption. The conclusion that public disclosure can help remedy misperceptions about the beliefs and behaviors of others is in the spirit of Jackson (2018), who finds that beliefs and aggregate actions can be corrected by public disclosure when there are positive strategic complementarities in the actions of different agents. Our findings differ in two key ways.

First, our finding in the setting with smart agents is not based on network asymmetry. Even if all agents are equally well-connected, disclosure helps correct misperceptions about others. In contrast, network asymmetry (and its consequence, the friendship paradox) is required in Jackson’s model for disclosure to have an effect.

Second, our model is not based upon strategic complementarities, which drive the effects in Jackson’s model. The driving force behind our results is naivete about visibility bias.\(^{28}\) So whereas Jackson’s approach is about the interaction between naivete about the friendship paradox and payoff/preference complementarities, our approach is focuses solely on belief biases. So even in our network model, in which the friendship paradox is involved, the mechanism by which it operates to amplify overconsumption (which is present even without it) is quite different.

## 5 Extensions

Both to address the generality of our conclusions and to address interesting additional issues, we now consider extensions of the base model which, for tractability, make some stronger assumptions.

First, in reality people differ in age, and someone who has saved heavily when young will have more resources for consuming when old. This raises the question of whether

\(^{28}\)This can be relevant for empirical applications, since some strategic complementarities for some consumption activities are likely to be modest or even negative. For example, congestion effects can induce negative strategic complementarity. If others consume heavily by buying cars, the resulting traffic makes it less desirable to own a car.
overestimation of the consumption of the old might lead to an inference of high past saving, where that saving would reflect information suggesting a high rather than low risk of wealth disaster. We address this topic in Subsection 5.1.

Second, in reality people differ in their wealths, which changes the learning problem because observed consumption in many bins is an indication that the targets of observation have high wealth, not just favorable signals about the probability of no wealth disaster. This leads to empirical implications about wealth dispersion and overconsumption in Subsection 5.2.

Third, we have so far assumed a fixed riskfree interest rate. We allow for increasing supply of debt as a function of the interest rate in Subsection 5.3. 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 similar results as in the base model arise under some technical variations. We consider other utility functions in Appendix B. We then depart from the assumption that the maximal fraction of full consumption bins is 100%, in Appendix C. Specifically, our base model made the assumption that when the agent is maximally optimistic, and therefore consumes $\frac{W}{2}$ at date 1, that this involves consumption in all bins. Our extension allows for consumption of $\frac{W}{2}$ to leave some bins empty. This variation is also useful for the analysis in Sections 5.2 and 5.3.

For the variations we study in this section, we make the additional assumption that the number of prior and consumption observations, $Q$ and $M$, are very large, so that agents’ priors are very close to $p$ and $z_n$ is very close to $E[z]$. The fraction $\xi = \frac{M}{Q}$ is still an arbitrary positive number. We may think of this as studying the limit of a sequence of economies as $Q \to \infty$, with $M = \xi Q$ in each economy.

5.1 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. In practice, young people sometimes observe the consumption of older people who are consuming from their savings. What inference does a young agent draw if old agents’ consumption seems to be unexpectedly high? One possibility is that old agents saved a lot when they were young, because they viewed the risk of a wealth disaster as high. Alternatively, it could be that in the current period many old agents have had favorable wealth realizations (no disaster), and therefore have high resources available for current consumption. These effects promote opposite inferences, so a priori it is unclear which effect dominates.

To address this issue, 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. Specifically, the fraction $\lambda \in [0, 1]$ of the bin observations are of the young, and the
remaining fraction $1 - \lambda$ of observations is of the old. 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 lower $\lambda$. On the other hand, the old may act as role models for the young, leading to higher $\lambda$. In addition to bias toward observing young or old, $\lambda$ 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, then since all agents live exactly two dates, we can think of $\lambda$ as a summary statistic for population growth. The population pyramid will be such that $\lambda$ is high in rapidly growing populations.

Young agents observe a random sample of consumption from each cohort, i.e., $\lambda M$ observations are from the young generation, and $(1 - \lambda) M$ from the old. For each observation, they know which type they are observing. Visibility bias, as previously specified in (8), is present for both type of observations.

We study stationary equilibrium in which the average estimated probability of no wealth disaster, $\bar{\mathbb{p}}$, is constant over time. Moreover, we assume that the $\epsilon$-shock is independent across agents (though still identical in distribution), to avoid systematic variations in aggregate consumption across time. 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 $2p - \bar{p} < p$. Without visibility bias, by reasoning similar to that leading to (9) equilibrium average beliefs satisfy

$$\bar{p} = p + \xi \left[ \lambda \bar{p} + (1 - \lambda)(2p - \bar{p}) \right] \left( 1 + \xi \right),$$

which has the unique solution $\bar{p} = p$. In this equilibrium, agents in both cohorts consume on average $pW/2$, observations of young and old consumption are consequently equally informative, and young agents update accordingly.

Intuitively, on the RHS of (25), for given $p$, higher $\bar{p}$ is associated with higher consumption by the young, which increases the average inference from observation of the bins of the young (as reflected in the first term within the brackets). But higher $\bar{p}$ also reduces the consumption of the old, which contributes negatively to the inference (the second term within the brackets). These effects balance when $\bar{p} = p$.

When there is visibility bias, observation of the bins of the young and the old are biased toward full bins, as reflected in the $S_r$ function, so the equilibrium condition (25) is replaced by

$$\bar{p} = p + \xi S_r \left[ \lambda \bar{p} + (1 - \lambda)(2p - \bar{p}) \right] \left( 1 + \xi \right).$$

33
This is the OLG extension of the model with visibility bias. The expression reduces to (25) when \( \tau = 1 \).

**Proposition 9** In the OLG extension of the model with visibility bias, there is a stationary equilibrium satisfying the following properties:

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

2. The equilibrium probability estimate, \( \bar{p} \) is increasing in the fraction of young agents, \( \lambda \).

3. When \( \lambda = 0 \),

   \[
   \bar{p} = \frac{1 + \xi(\tau + 1) + p(3 + 2\xi)(\tau - 1) - \sqrt{V}}{2(1 + \xi)(\tau - 1)},
   \]

   where

   \[
   V = (1 + \xi(\tau + 1) + p(3 + 2\xi)(\tau - 1))^2 - 4p(1 + \xi)(\tau - 1)(1 + 2p(\tau - 1) + 2\xi\tau). \tag{28}
   \]

An implication of Proposition 9 is that we expect overconsumption to be more severe in economies with rapid population growth.

Figure 2 shows the equilibrium probability estimate, \( \bar{p} \) as a function of the true probability, \( p \), for the limiting cases \( \lambda = 0 \) (given by (27-28)) and \( \lambda = 1 \) (given by (13-14)). When \( 0 < \lambda < 1 \), the equilibrium lies between the two increasing concave curves, and is increasing in \( \lambda \) for a fixed \( p \).

At the start of this subsection we suggested that the inference drawn from observing the old is potentially mixed, but it turns out that the effect in equilibrium is unambiguous. As a reminder, the argument for why the effect might be mixed is that, owing to visibility bias, observers may think they see the old consuming heavily, potentially implying high past saving. Such high saving would be indicative of information suggesting a high need to save.

To understand why, nevertheless, there must be overconsumption in equilibrium, recall that optimally agents split in half their expected wealth (net of disasters) between the two dates. So it is when \( p \) is high that, in an equilibrium with no visibility bias, average consumption of the old is high. (Recall that all agents think they are living in a world without visibility bias.) Owing to visibility bias, the young think they see evidence of heavy consumption by others, and they believe that the average consumption levels of the young and old are equal. It follows that the young observers think that \( p \) is high; it is high \( p \) that generates high \( \bar{p} \) and high average consumption by both young and old.
Figure 2: Equilibrium overconsumption in OLG model with visibility bias. The figure shows equilibrium without overconsumption (\(\bar{p} = p\); Parameter: \(\tau = 1\)), with overconsumption and large fraction of old (\(\lambda = 0\)), and with overconsumption and large fraction of young (\(\lambda = 1\)). When \(0 < \lambda < 1\), the equilibrium lies between \(\lambda = 0\) and \(\lambda = 1\), and is increasing in \(\lambda\). Parameters: \(\tau = 2\), \(\xi = 1\).

For example, consider the special case in which \(\lambda = 0\) so that observers only see the old. Owing to visibility bias, young observers conclude that the old are consuming heavily, which implies a higher-than-actual value of \(p\). In other words, the young conclude that the old are consuming heavily now, and did so equally when young (low past saving). Based on their average overestimate of \(p\), the young overconsume.

This intuition makes clear why, when there is visibility bias, the two effects described at the start of this section do not offset. As argued there, when a young agent sees evidence suggesting that an old agent is consuming heavily, the observer might conclude either that the old agent previously saved heavily, indicating high anticipation of disaster; or that the old agent was not hit by a disaster, which is favorable information. What explicit modeling reveals is that these effects do not offset, because observers understand that the old optimized when they were young, and think that the old did so rationally. So the perception of high average consumption of the old is taken as a favorable indication that the old, when young, were foresaw high expected wealth to allocate across time periods.

This intuition also makes clear why in equilibrium, overconsumption is greater when observation is more heavily tilted toward the young (\(\lambda\) 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 $\lambda$ (sampling from the young) leads to more favorable inferences about $p$, and therefore greater overconsumption.

These findings suggest that salient public disclosure of the consumption of the old can help address the problem of overconsumption. This intervention differs from that considered earlier, as it involves disclosing the average consumption of a subset of the population, not of the entire population. The effect of such disclosure is effectively to push the model in the direction of low $\lambda$ (in which there is more observation of consumption of the old). As we have shown, lower $\lambda$ decreases aggregate consumption. It is interesting that even when disclosing aggregate consumption does not help, disclosing the consumption of the right subset of the population does help— but does not fully remedy the problem. At best it only reduces consumption to that of the $\lambda = 0$ case.

5.2 Information Asymmetry about Wealth

So far, we have assumed that all agents have the same initial wealth, $W$. We now generalize to allow for wealth dispersion in the population, and ignorance of the wealths of others. Intuitively, the inference an observer draws about the information of others based on observation of another’s consumption is weaker if the observer does not know the target’s wealth, because high apparent consumption of a target of observation could come either from the target possessing a favorable signal (indicating low risk of disaster), or from possessing high wealth. Observers will therefore not, on average, revise their estimate of $p$ upward as aggressively as they do when there is no wealth dispersion. Wealth dispersion therefore reduces equilibrium overconsumption. This contrasts sharply with the Veblen wealth-signaling approach, in which it is precisely the fact that there is uncertainty about wealth that causes overconsumption to serve as a signal.

To explore the effects of wealth dispersion, we study an economy in which the consumption fraction is $f < 1$, in the following sense. So far, we have assumed that when the investor consumes half of potential wealth at date 0, $c_0 = W/2$, that all consumption bins are full. We now instead assume that in this circumstance, only a fraction $0 < f \leq 1$ of the bins are full. Moreover, a fraction $\lambda$ of the population has wealth $(1 + \Delta)W$ (the wealthy fraction), where $\Delta > 0$, and $(1 + \Delta)f < 1$, a fraction $\lambda$ has wealth $(1 - \Delta)W$ (the poor fraction), and the remaining fraction $1 - 2\lambda$ has wealth $W$ (the medium fraction). The average wealth is then still $W$, but the higher $\Delta$ is, the higher the economy’s wealth dispersion.

For this analysis, we need to assume that the consumption fraction is less than one, as otherwise the wealthy group would potentially consume so heavily that more than 100% of the consumption bins are full. We therefore assume that $\Delta$ is not too large. Specifically,
we assume that
\[ \Delta \leq \frac{1}{1 + \frac{2}{(\tau - 1)}}. \]  

(29)

For large \( \tau \), this restriction is very weak, basically implying that no agent has negative wealth or wealth very close to zero.

Agents know the economy’s wealth distribution (\( \lambda \) and \( \Delta \)), and for simplicity we assume that all their consumption bin observations come from one or more agents of the same wealth type. An agent who observes the consumption fraction \( z \) then forms his posterior beliefs about \( p \), taking into account that an observation of high consumption could reflect high wealth, not just optimistic beliefs, on the part of the target of observation.

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

**Proposition 10**  The equilibrium probability estimate \( \bar{p} \) is decreasing in wealth dispersion, \( \partial \bar{p}/\partial \Delta < 0 \), as is the overconsumption factor.

Proposition 10 predicts that savings rates increase with wealth dispersion. This is the opposite of what is expected based upon Veblen wealth-signaling considerations. 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). Greater information asymmetry about wealth intensifies the effect, by increasing the potential improvement in wealth perceptions that can be achieved by signaling. This is reflected, for example, in a comparative statics of Charles, Hurst, and Roussanov (2009) in which a parameter shift that increases wealth dispersion results in greater consumption signaling. More generally, in Veblen-style models wealth signaling through consumption vanishes when wealth dispersion is zero. This implies, moving up from zero wealth dispersion, an average tendency for greater wealth dispersion to induce greater overconsumption, though not necessarily monotonically.\(^{29}\)

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 the visibility bias approach, in contrast with the implication of wealth-

\(^{29}\)Consistent with this idea, Charles, Hurst, and Roussanov (2009) find empirically that greater dispersion of reference group income is associated with higher visible spending for minorities, consistent with our theory. On the other hand, they report that higher dispersion in reference group income significantly lowers White visible spending.
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.\textsuperscript{30}

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 10 counterfactually implies a rising savings rate. However, all else was not equal. From the standpoint of our model, a more fundamental effect (which holds even in our base model) comes from the dramatic transformation of electronic communications and social networks. As discussed in Section 3.2, 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.\textsuperscript{31}

As discussed in the introduction, agents who disagree may overconsume owing to the expectation of profiting by trading against others Heyerdahl-Larsen and Walden (2017). Although our model allows for difference in beliefs, our overconsumption finding is not driven by this speculative effect. Agents do not trade in speculative claims based on the occurrence of wealth disasters.\textsuperscript{32}

The speculative trading theory shares with our model the prediction that high wealth dispersion reduces overconsumption. However, the models are empirically distinguishable in various ways. Our model does not share predictions of the speculative trading theory that relate individual or aggregate consumption to proxies for disagreement (such as trading volume or derivatives market open interest) or to individual participation in speculative markets. Our model also has various distinctive implications about how social network connections affect overconsumption, and the effects of public disclosure of the consumption or saving of others.

\textsuperscript{30}Concern for relative consumption, as in ‘keeping up with the Joneses’, can induce a fear of falling behind which raises precautionary savings (Harbaugh 1996). We are not aware of any results in the standard keeping-up-with-Joneses model relating overconsumption to wealth dispersion (especially holding constant the average level of wealth). Bertrand and Morse (2016) propose that a variant of the keeping-up-with-Joneses approach suggests a negative relationship between income inequality and the savings rate of non-rich households.

\textsuperscript{31}Furthermore, it can be argued that the relevant time series shift in wealth dispersion is downward rather than upward. Our result that wealth dispersion reduces overconsumption derives from asymmetric information—the unobservability of others’ wealths rather than dispersion per se. In some countries, the rise of the internet has made observation of others’ wealths or incomes easier than in the past through search of government or other archives. To the extent that this is true, our approach implies greater overconsumption, because people attribute their high observations of others’ consumption to favorable information possessed by others rather than to high wealths. This effect reinforces the other time series shifts we’ve described, implying a shift over time toward greater overconsumption.

\textsuperscript{32}Even if wealth disasters were systematic, apart from random variations (which can be very small if $Q$ and $M$ are large), agents have essentially the same (wrong) beliefs about $p$, and will therefore not bet against each other. Neither will they buy insurance against $\epsilon$ shocks at the unbiased market value.
5.3 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.

The base model is highly tractable, but 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 make further adjustments to the model to prevent this possibility.

As in the base model, agent utility is defined by

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

(30)

where we focus on the case \( \alpha = 1/2 \). Given a one-period interest rate of \( r \), an agent’s budget constraint is now

\[ c_1 = (1 + r)(W - c_0) - \epsilon, \]

(31)

For tractability, in this section we assume a less severe bad outcome than in the base model, so that time-0 consumption remains nonnegative for a larger range of interest rates. We therefore assume that

\[ \epsilon = \begin{cases} 
0 & \text{with probability } p \\
\frac{W}{2} & \text{with probability } 1 - p,
\end{cases} \]  

(32)

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

\[ \frac{1 - r + 2r^2 + (1 + r)\hat{p}}{4 + 4r + 2r^2} = g + f\hat{p}, \]  

(33)

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

\[ \bar{c}_0 = g + f\bar{p}. \]  

(34)

We assume that there is an equally large set of investors who are not exposed to \( \epsilon \) risk, i.e., for whom \( c_1 = (1 + r)(1 - c_0) \). 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. Since these outsiders have no disaster risk, their consumption does not depend on their inferences about \( p \). Their role in the model is to provide someone for individual investors to trade with, so that even in our exchange economy, beliefs can
affect the equilibrium per capita consumption of the individual investors, rather than just
inducing an adjust to the interest rate. In other words, if they become overoptimistic, they
can potentially borrow more and overconsume.\textsuperscript{33}

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 time-0 consumption
of institutions, given $r$, is

$$
\bar{c}_I^0 = \frac{1 + r^2}{2 + 2r + r^2}.
$$

(35)

We assume free disposal of the consumption good, so the equilibrium interest rates 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 $r \leq 1/2$.\textsuperscript{34}

The aggregate endowed quantity of the time-0 consumption good per individual investor,
$\bar{c}^e$ is fixed. 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

$$
\bar{c}^e = \bar{c}_I^0 + \bar{c}_0,
$$

(36)

where the terms on the right are functions of $r$. By (34,35), and by our assumption the
endowment is such that when $r = 0$ a market with unbiased investors would clear, we have

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

(37)

so the general market clearing condition becomes

$$
\bar{c}_I^0 + \bar{c}_0 = \frac{1}{2} + \frac{1 + p}{4}.
$$

(38)

The LHS is the aggregate per capita endowed consumption.

If, owing to visibility bias, $\bar{p} \neq p$, market clearing implies that $r$ must adjust so that
aggregate demand is unchanged, i.e., is equal to the endowment, so in equilibrium

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

(39)

\textsuperscript{33}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.

\textsuperscript{34}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.
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 contain consumption is given by (34). Since agents suffer from visibility bias, they observe the fraction $S_{\tau}(g + f\bar{p})$. By Bayes rule, the agents then arrive at the posterior probability estimate

$$\hat{p} = R(p, z, \xi, f, g),$$  \hspace{1cm} (40)

given that the fraction of bins that agents observe are full is $z$, where the function $R$ is defined in the appendix (see the proof of Proposition 11). An equilibrium is then an outcome in which markets clear, so that (34) holds, and agents’ posterior beliefs are in line with their biased observations, $\bar{p} = R(p, S_{\tau}(g + f\bar{p}), \xi, f, g)$. For tractability, we focus on the case when $\xi = 1$ (so that agents put comparable weight on prior information and on social observations).

It is now possible to show the existence of equilibrium with the following properties:

**Proposition 11** *In equilibrium, under the above assumptions, the equilibrium probability estimate, overconsumption, and the interest rate, are all increasing in visibility bias, $\tau$.*

### 6 Concluding Remarks

We examine how social influence 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 (and a well-known manifestation of it, the availability heuristic), people infer that low saving is a good idea. 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 arbitrarily strong. The effects in the model can also bring about pluralistic ignorance about the savings rates of others, wherein people think that others are consuming even more heavily than they really are.

The visibility bias approach offers a simple explanation for one of the most puzzling and important stylized facts about 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.

In contrast with the present bias (hyperbolic discounting) theory of overconsumption, the effects here are induced by social observation and interaction. Our approach can therefore be distinguished from present bias using proxies for sociability and observability, such as urban versus rural, and survey responses about sociability (see, e.g., Hong, Kubik, and Stein (2004), Brown et al. (2008), Christelis, Georgarakos, and Haliassos (2011), and Georgarakos and Pasini (2011)). Our approach also differs in offering predictions about how population-level characteristics such as wealth variance affect consumption.

Also, our visibility bias approach is not based on a link between an agent’s participation in speculative markets and his overconsumption, in contrast to overconsumption associated with heterogeneous beliefs as in Heyerdahl-Larsen and Walden (2017).

The effect of wealth dispersion in our model contrasts with the implications of the Veblen and keeping-up-with-the-Joneses approaches. In at least some versions of the latter approach, wealth dispersion encourages the poor to consume more in emulation of the wealthy. The Veblen wealth-signaling approach broadly implies that a comparative statics shift from certainty to information asymmetry about others’ wealths (i.e., a rise in wealth dispersion) implies greater overconsumption. In our setting greater information asymmetry dilutes the inference from high observed consumption that others have favorable information about the risk of a wealth disaster. In consequence, equilibrium consumption is lower, the opposite prediction. The visibility bias approach also helps explain high variation in savings rates across countries and ethnic groups, because even modest differences in beliefs can be amplified through social influence.

In contrast with the signaling, preference-based, and speculative disagreement approaches, our social learning bias approach implies that a relatively simple policy intervention can potentially increase saving. The relevant intervention is to provide—in highly salient form—accurate information about how much peers save, or their attitudes toward consumption. Indeed, there is evidence discussed earlier supporting this implication in specialized settings, such as the decisions of college students of how much to drink.

Advertising and media biases can further reinforce overconsumption for reasons very similar to those that we model. Advertisers have an obvious incentive to depict consumers using their products heavily. News media serve their clientele by highlighting interesting high-end products or consumption events (consider, e.g., the ‘Travel’ section of newspapers). These further contribute to the higher visibility of consumption than nonconsump-
tion. Of course, there is advertising of financial saving vehicles as well. But it is much easier to vividly depict individuals consuming heavily at restaurants or exotic locations than to depict individuals saving heavily.

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 substantial cyclical shifts in overconsumption. Such fluctuations may help explain local consumption booms and aggregate business cycles. This might potentially provide an interesting contrast to Keynesian ideas about business cycles deriving from resource underutilization and underconsumption.
Appendices

A Proofs

Proof of Proposition 1: Part 1 follows from noting that (12,13) implies that \( \bar{p} = p \) if and only if \( -4(1-p)p(\tau-1)2\xi(1+\xi) = 0 \), 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 \( \tau \), and taking the derivative w.r.t. \( x \), leading to \( \frac{\partial \bar{p}}{\partial x} = \left( x + p - \xi + 2p\xi - \sqrt{(p-x+\xi)^2 + 4px(1+\xi)} \right) r(x) \), where \( r(x) > 0 \). It then follows from the fact that \( (x + p - \xi + 2p\xi)^2 - ((p-x+\xi)^2 + 4px(1+\xi)) = -4(1-p)p\xi(1+\xi) < 0 \), that \( \frac{\partial \bar{p}}{\partial x} < 0 \), and thus \( \frac{\partial \bar{p}}{\partial \tau} > 0 \). Parts 2 and 4 follow immediately by taking the limit of (12,13) as \( \xi \) and \( \tau \) become large. To show Part 3, note that \( \bar{p} \) can be written as

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

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

Proof of Proposition 3: Equilibrium with the public noisy signal is defined via the relation:

\[
\bar{p}^V = \frac{p + \xi S_\tau(\bar{p}) + \alpha \bar{p}}{1 + \xi + \alpha}, \quad \bar{p} = \phi p + (1-\phi)p^V,
\]

which is equivalent to

\[
\bar{p}^V = \frac{p + \xi S_\tau(\bar{p})}{1 + \xi}, \quad \hat{\xi} = 1 + \alpha \phi,
\]

which we in turn rewrite as

\[
\hat{\xi} = \frac{\bar{p}^V - p}{S_\tau(\bar{p}) - \bar{p}^V}.
\]

It follows that

\[
\frac{d\hat{\xi}}{dp^V} = \frac{1}{S_\tau(\bar{p}) - \bar{p}^V} + \frac{1 - (1-\phi)S'_\tau(\bar{p})}{(S_\tau(\bar{p}) - \bar{p}^V)^2} > 0,
\]

(41)
since $S_r(\bar{p}) > \bar{p}'$ and $S'_r(\bar{p}) < 1$, and thus that
\[
\frac{d\bar{p}'}{d\xi} > 0.
\]
Since $\frac{d\xi}{da} = -\frac{\phi}{(1+\alpha\phi^2)} < 0$, the result follows.

**Proof of Proposition 4:** Denote by $\bar{p}$ the equilibrium probability estimate in (14) and by $\bar{p}'$ the equilibrium in (18). It follows from these equations that
\[
\bar{p}' - \bar{p} = \frac{\xi}{1 + \xi} (S_r(\bar{p}') - S_r(\bar{p})) + \frac{\alpha}{1 + \xi} (1 - \bar{p}' - S_r(1 - \bar{p}')).
\]
Define the function
\[
R(x) = x - \bar{p} - \left(\frac{\xi}{1 + \xi} (S_r(x) - S_r(\bar{p})) + \frac{\alpha}{1 + \xi} (1 - x - S_r(1 - x))\right),
\]
and note that $R(\bar{p}') = 0$, and that $R$ is a continuous function of $x \in [0,1]$. From (15), it follows that $\bar{p} > \frac{\xi}{1 + \xi} S_r(\bar{p})$, and therefore, since $S_r(0) = 0$, $S_r(1) = 1$, that $R(0) < 0$. Moreover, since $S_r(y) > y$, for $y \in (0,1)$, it follows that $R(\bar{p}) = \frac{\alpha}{1 + \xi} (S_r(1 - \bar{p}) - (1 - \bar{p})) > 0$. 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 5:** 1. Define the mapping $F : \mathbb{R}_+^N \to \mathbb{R}_+^N$ by
\[
(F(w))_n = \frac{p + d_n \xi S \left(\frac{1}{d_n} \sum_{m \in D_n} w_m\right)}{1 + d_n \xi}, \quad n = 1, \ldots, N.
\]
An equilibrium is then a fixed point to this mapping, $\bar{p} = F(\bar{p})$. It is easy to see that $F$ is nondecreasing in each of its arguments: $w^2 \geq w^1 \to F(w^2) \geq F(w^1)$, that $F(p, p, \ldots, p) = \left(\frac{p+\xi S(p)}{1+\xi}, \ldots, \frac{p+\xi S(p)}{1+\xi}\right)' \equiv z \geq (p, p, \ldots, p)'$, and that $F(1,1,\ldots,1) = \left(\frac{p+\xi S(1)}{1+\xi}, \frac{p+\xi S(1)}{1+\xi}, \ldots, \frac{p+\xi S(1)}{1+\xi}\right)' \equiv y \leq (1,1,\ldots,1)'$. It follows that $F$ maps the convex compact set $S \equiv [z,y]^N$ into itself. Since $F$ is a continuous mapping, Brouwer’s theorem implies that $F$ has a fixed point, i.e., that there exists an equilibrium in $S$, and since $S \subset [p,1]^N$, the existence result follows.

For $\tau = 1$, (20) reduces to the linear algebraic equation
\[
\bar{p} = (I + \xi \text{diag}(d_1, \ldots, d_N))^{-1} (p1 + \xi E\bar{p}),
\]
or equivalently,
\[(I + \xi \text{diag}(d_1, \ldots, d_N) - \xi E)\tilde{p} = p1.\]

Here, the \(1\) is a vector of ones, \(1 = (1, \ldots, 1) \in \mathbb{R}^N\). It is easy to verify that \(\tilde{p} = p1\) is a solution. Since the matrix \(A \overset{\text{def}}{=} (I + \xi \text{diag}(d_1, \ldots, d_N) - \xi E)\) is diagonally dominant \((A_{nn} - \sum_{m \neq n} |A_{mn}| = 1 > 0, n = 1, \ldots, N)\), it is invertible, so this solution is unique when \(\tau = 1\). When \(\tau > 1\), it also immediately follows that \(p1\) is not a solution, since \(z \gg p1\) in this case, and the solution must lie in \(S = [z, y]^N\).

2. The case \(\tau = 1\) is already covered in part 1 of the proof, so w.l.o.g. assume that \(\tau > 1\). Consider the function \(G : (0, 1)^N \to \mathbb{R}\), defined by

\[
G(x) = \sum_{n=1}^{N} d_n R_n \log(x_n) + \frac{1}{2} \sum_{n,m=1}^{N} x_n E_{nm} x_m - \sum_{n=1}^{N} g_n x_n,
\]

\[
R_n = \frac{d_n \xi \tau}{(1 + d_n \xi)(\tau - 1)^2},
\]

\[
g_n = \frac{d_n}{\tau - 1} + \sum_{m \in \mathcal{D}_n} f_m,
\]

\[
f_n = \frac{p}{1 + d_n \xi} + \frac{d_n \xi \tau}{(1 + d_n \xi)(\tau - 1)}.
\]

The gradient of \(G\) is \(\nabla G(x) \in \mathbb{R}^N = a + Ex - g\), where \(a = (a_1, \ldots, a_N)'\), \(a_n = \frac{d_n R_n}{x_n}\), \((Ex)_n = \sum_{m \in \mathcal{D}_n} x_m\), and \(g = (g_1, \ldots, g_N)'\). A stationary point of \(G\) satisfies \(\nabla G(x) = 0\).

Defining the bijection \(\tilde{p} \leftrightarrow x\), via \(\tilde{p}_n \overset{\text{def}}{=} f_n - x_n\), it follows that at such a stationary point \(\frac{d_n R_n}{f_n - \tilde{p}_n} = g_n - \sum_{m \in \mathcal{D}_n} f_m + \sum_{m \in \mathcal{D}_n} \tilde{p}_m = \frac{d_n}{\tau - 1} + \sum_{m \in \mathcal{D}_n} \tilde{p}_m\), or equivalently

\[
\tilde{p}_n = f_n - \frac{\frac{d_n R_n}{\tau - 1} + \sum_{m \in \mathcal{D}_n} \tilde{p}_m}{\frac{d_n}{\tau - 1} + \sum_{m \in \mathcal{D}_n} \tilde{p}_m} = \frac{p}{1 + d_n \xi} + \frac{d_n \xi \tau}{(1 + d_n \xi)(\tau - 1)} - \frac{\frac{R_n}{\tau - 1} + \frac{1}{d_n} \sum_{m \in \mathcal{D}_n} \tilde{p}_m}{\frac{1}{\tau - 1} + \sum_{m \in \mathcal{D}_n} \tilde{p}_m}
\]

\[
= \frac{p}{1 + d_n \xi} + \frac{d_n \xi}{1 + d_n \xi} \left( \frac{1 - \frac{1}{\tau}}{(1 - \frac{1}{\tau})(\tau - 1)} \left( \frac{1}{\tau - 1} + \frac{1}{d_n} \sum_{m \in \mathcal{D}_n} \tilde{p}_m \right) \right)
\]

\[
= \frac{p}{1 + d_n \xi} + \frac{d_n \xi}{1 + d_n \xi} \left( \frac{(\tau - 1)(\frac{1}{\tau - 1} + \frac{1}{d_n} \sum_{m \in \mathcal{D}_n} \tilde{p}_m) - 1}{(1 - \frac{1}{\tau})(\tau - 1)} \left( \frac{1}{\tau - 1} + \frac{1}{d_n} \sum_{m \in \mathcal{D}_n} \tilde{p}_m \right) \right)
\]

or equivalently, \((F(\tilde{p}))_n\).
Thus, every equilibrium point, \( \bar{p} \), is equivalent to a stationary point of \( G \), \( x \), under the mapping \( \bar{p} \leftrightarrow x \), with

\[
x_n = \frac{p}{1+d_n \xi} + \frac{d_n \xi \tau}{(1+d_n \xi)(\tau-1)} - \bar{p}_n.
\]

It also follows that the set \( S \) under the \( \bar{p} \leftrightarrow x \) mapping corresponds to the set

\[
\{x\} \in U \overset{\text{def}}{=} \left[ 0, \frac{d_1 \xi}{1+d_1 \xi} \left( \frac{\tau}{\tau-1} - \frac{\tau p}{(\tau-1)p+1} \right) \right] \times \ldots \times \left[ 0, \frac{d_N \xi}{1+d_N \xi} \left( \frac{\tau}{\tau-1} - \frac{\tau p}{(\tau-1)p+1} \right) \right].
\]

Thus, if there is a unique stationary point of \( G \) in \( U \), then the corresponding equilibrium vector \( \bar{p} \) is unique in \( [p,1]^N \).

It is easy to check that the Hessian of \( G \), \( H_G(x) \in \mathbb{R}^{N \times N} \), has elements

\[
[H_G(x)]_{nm} = \begin{cases} 
-d_n R_n, & n = m, \\
E_{nm} & n \neq m.
\end{cases}
\]

It follows that for \( x_n \in \left[ 0, \frac{d_n \xi}{1+d_n \xi} \left( \frac{\tau}{\tau-1} - \frac{\tau p}{(\tau-1)p+1} \right) \right] \),

\[
[H_G(x)]_{nn} \leq -d_n \frac{R_n}{\left( \frac{d_n \xi}{1+d_n \xi} \left( \frac{\tau}{\tau-1} - \frac{\tau p}{(\tau-1)p+1} \right) \right)^2} = -d_n \frac{d_n \xi}{(1+d_n \xi)(\tau-1)^2} \left( \frac{\tau}{\tau-1} - \frac{\tau p}{(\tau-1)p+1} \right)^2 = -d_n \left( \frac{1+d_n \xi}{d_n \xi} \right) \left( \frac{1}{\tau} \right) ((\tau-1)p+1)^2.
\]

Since \( \sum_m E_{nm} = d_n \) for all \( n \), it follows that if

\[
\left( \frac{1+D \xi}{D \xi} \right) \left( \frac{1}{\tau} \right) ((\tau-1)p+1)^2 > 1,
\]

then \( H_G \) is diagonally dominant with negative diagonal elements, in the whole of \( U \) and thus the Hessian is negative definite in this region. Standard theory of optimization then in turn implies that a stationary point of \( G \) is unique in \( U \), and thus that \( \bar{p} \) is unique in \( S \), and therefore also in \( [p,1]^N \). The condition (49) is obviously equivalent to (21).

3. The result follows immediately from the fact that when \( \tau \to \infty \), both \( z_n \) and \( y_n \), as defined in part 1 of the proposition, converge to \( \frac{p + d_n \xi}{1+d_n \xi} \) for all \( n \).
Proof of Proposition 6: Conjecture an equilibrium in which \( \bar{p}_1 = \bar{p}_2 = \cdots = \bar{p}_N = w \), since \( \frac{1}{d_n} \sum_{m \in D_n} \bar{p}_m = w \) for all \( n \), (20) reduces to the condition \( \bar{p}_n = \frac{p + d \xi S_\tau(\bar{p}_n)}{1 + d \xi} \), which has solution \( \bar{p}_n = B(\tau, p, d \xi), n = 1, \ldots, N \). \( \square \)

Proof of Proposition 7: As evident from Proposition 5, equilibrium consumption satisfies \( \bar{p}^P > p, \bar{p}^C > p \). Assume that \( \bar{p}^P \geq \bar{p}^C \). It follows that \( \bar{p}^C \leq \zeta \leq \bar{p}^P \), and consequently that \( S_\tau(\zeta) \geq S_\tau(\bar{p}^C) \). Consequently (since \( d \xi + d P > 1 \)),

\[
\frac{p + (dC + DP) \xi S_\tau(\zeta)}{1 + (dC + DP) \xi} > \frac{p + \xi S_\tau(\bar{p}^C)}{1 + \xi},
\]

and thus \( \bar{p}^C > \bar{p}^P \), leading to a contradiction. It follows that \( \bar{p}^C > \bar{p}^P \).

For the second part of the theorem, note that it follows from (22, 23) that in equilibrium \( F(\bar{p}^C, d^C) = 0 \), where

\[
F(\bar{p}^C, d^C) = p + (dC + DP) \xi S_\tau \left( \frac{dC}{dC + DP} \bar{p}^C + \frac{dP}{dC + DP} \left( \frac{p + \xi S_\tau(\bar{p}^C)}{1 + \xi} \right) \right) - \bar{p}^C (1 + (dC + DP) \xi)
\]

where \( d = dC + DP \). Now,

\[
\frac{\partial F}{\partial \bar{p}^C} = d \xi S'_\tau(\zeta) \left( \frac{dC}{dC + DP} + \frac{dP}{dC + DP} \frac{\xi}{1 + \xi} S'_\tau(\bar{p}^C) \right) - (1 + d \xi)
\]

\[
\leq d \xi - (1 + d \xi)
\]

\[
< 0,
\]

where the inequality follows from the fact that \( \bar{p}^C > \frac{1}{2}, \zeta > \frac{1}{2} \), and \( S'_\tau(x) < 1 \) when \( x > \frac{1}{2} \).

Moreover,

\[
\frac{\partial F}{\partial dC} = \xi S_\tau(\zeta) + d \xi S'_\tau(\zeta) \frac{d \xi}{dC} - \bar{p}^C \xi
\]

\[
> 0
\]

which follows from the fact that \( S_\tau(\zeta) \geq \bar{p}^C \) (which in turn follows from (22)), and that \( \frac{d \xi}{dC} = \frac{1}{\bar{p}^C} (1 - \bar{p}^D) > 0 \). Therefore, by the inverse function theorem, \( \frac{d \xi}{dC} = -\frac{\bar{p}^C}{\bar{p}^D} > 0 \).

Finally, from (23) it follows that \( \bar{p}^D \) is strictly increasing in \( \bar{p}^C \). Both \( \bar{p}^C \) and \( \bar{p}^D \) are therefore increasing in \( dC \), as is then aggregate consumption. We are done. \( \square \)
Proof of Proposition 8: For a symmetric network, equation (24) reduces to
\[ \bar{p} = \frac{p + d\xi S_{\tau}(\bar{p}) + d\alpha \bar{p}}{1 + d\xi + d\alpha}. \] (50)
since \( \bar{p}_n = \bar{p} \) and \( d_n = d \) for all agents. It follows immediately that any solution to (50) equivalently satisfies
\[ \bar{p} = \frac{p + d\xi S_{\tau}(\bar{p})}{1 + d\xi}, \] (51)
i.e., is also an equilibrium in the economy without public signal. \( \blacksquare \)

Proof of Proposition 9: It is easy to verify that the solution to the equilibrium condition (26) is
\[ \bar{p} = \frac{-1}{2(2\lambda - 1)(1 + \xi)(\tau - 1)} \left( 1 - 3p + 4\lambda p + \xi - 2p\xi + 2\lambda p\xi + 3p\tau - 4\lambda p\tau + \xi\tau + 2p\xi\tau - 2\lambda p\xi\tau - \sqrt{V} \right), \] (52)
where
\[ V = -4(2\lambda - 1)p(1 + \xi)(\tau - 1) \left( -1 + 2(\lambda - 1)p(\tau - 1) + 2(\lambda - 1)\xi\tau \right) + \left( 1 - p(-3 - 2\xi + 2\lambda(2 + \xi)(\tau - 1) + \xi(1 + \tau - 2\lambda\tau) \right)^2. \]
It is also easy to verify that the solution reduces to (27,28) when \( \lambda = 0 \), and to (12,13) 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 \( \lambda \), except possibly at \( \lambda = 1/2 \).

We next define \( x = 2\lambda - 1 \in [-1,1] \), and rewrite (52) as
\[ \bar{p} = a + \frac{b(\sqrt{1 - cx + dx^2} - 1)}{2x(\tau - 1)}, \]
where
\[ a = \frac{p(2 + \xi)(\tau - 1) + \xi\tau}{2(1 + \xi)(\tau - 1)}, \]
\[ b = 1 + p(\tau - 1), \]
\[ c = \frac{2\xi(-p(\tau - 1)^2 + p^2(\tau - 1)^2 + \tau)}{(1 + \xi)(1 + p(\tau - 1))^2}, \]
\[ d = \frac{\xi^2(p(1 - \tau) + \tau)^2}{(1 + \xi)^2(1 + p(\tau - 1))^2}. \]
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 $\lambda$ at $\lambda = 1/2$ too. Thus, since $\bar{p}^{\lambda=1} > p$, $\bar{p}^{\lambda}$ depends continuously on $\lambda$, 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
\[
\frac{d\bar{p}}{dx} = \frac{b}{4(\tau - 1)} \times \frac{\sqrt{1-cx+dx^2 + \frac{c}{2}x}}{\sqrt{1-cx+dx^2}},
\]
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
\[
4d^2 - c^2 = \frac{16(1-p)p\xi^2(\tau-1)^2\tau}{(1+\xi)^2(1+p(\tau-1))^2} > 0,
\]
so the condition is indeed satisfied.

**Proof of Proposition 10:**

We first state and prove the following lemma, which characterizes the equilibrium probability estimate:

**Lemma A.1** The equilibrium probability estimate is the solution to the equation
\[
\bar{p} = qR(p, S_r(f\bar{p}(1-\Delta)), \xi, f(1+\Delta)) + (1-2q)R(p, S_r(f\bar{p}), \xi, f(1+\Delta)) + qR(p, S_r(f\bar{p}(1+\Delta)), \xi, f(1+\Delta)),
\]

where the function $R$ is defined in (61).

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 $f\bar{p}(1+\Delta)$ of bins rather than the average, $f\bar{p}$). 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.1: For \( \alpha, \beta, f_1, f_2 \in (0, 1] \), \( \xi > 0 \), define

\[
L(\alpha, \beta, \xi, f_1, f_2) = \lim_{Q \to \infty} \frac{J(\alpha, \beta, \xi, f_1, Q, 0)}{J(\alpha, \beta, \xi, f_2, Q, 0)},
\]

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

\[
L(\alpha, \beta, \xi, f_1, f_2) = \begin{cases} 
\infty, & |f_1 - \frac{\beta}{\alpha}| < |f_2 - \frac{\beta}{\alpha}|, \\
0, & |f_1 - \frac{\beta}{\alpha}| > |f_2 - \frac{\beta}{\alpha}|.
\end{cases}
\]

(54)

An agent with prior \( \tilde{p} \), who observes a fraction \( \beta \) 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 \((q, 1 - 2q, q)\) who consume the fraction \((f(1 - \Delta), f, f(1 + \Delta))\) of the bins, respectively, will update—using Bayes rule—to the posterior:

\[
\hat{p} = \frac{qJ(\alpha, \beta, \xi, (1 - \Delta)f, Q, 1) + (1 - 2q)J(\alpha, \beta, \xi, f, Q, 1) + qJ(\alpha, \beta, \xi, (1 - \Delta)f, Q, 1)}{qJ(\alpha, \beta, \xi, (1 + \Delta)f, Q, 0) + (1 - 2q)J(\alpha, \beta, \xi, f, Q, 0) + qJ(\alpha, \beta, \xi, (1 + \Delta)f, Q, 0)} \cdot g_1 + \frac{J(\alpha, \beta, \xi, (1 - \Delta)f, Q, 1)}{J(\alpha, \beta, \xi, (1 - \Delta)f, Q, 0)} g_2 + \frac{J(\alpha, \beta, \xi, (1 + \Delta)f, Q, 1)}{J(\alpha, \beta, \xi, (1 + \Delta)f, Q, 0)} g_3,
\]

where

\[
g_1 = \frac{qJ(\alpha, \beta, \xi, (1 - \Delta)f, Q, 0) + (1 - 2q)J(\alpha, \beta, \xi, f, Q, 0) + qJ(\alpha, \beta, \xi, (1 + \Delta)f, Q, 0)}{(1 - 2q)J(\alpha, \beta, \xi, f, Q, 0)},
\]

\[
g_2 = \frac{qJ(\alpha, \beta, \xi, (1 - \Delta)f, Q, 0) + (1 - 2q)J(\alpha, \beta, \xi, f, Q, 0) + qJ(\alpha, \beta, \xi, (1 + \Delta)f, Q, 0)}{qJ(\alpha, \beta, \xi, (1 + \Delta)f, Q, 0)},
\]

\[
g_3 = \frac{qJ(\alpha, \beta, \xi, (1 - \Delta)f, Q, 0) + (1 - 2q)J(\alpha, \beta, \xi, f, Q, 0) + qJ(\alpha, \beta, \xi, (1 + \Delta)f, Q, 0)}{qJ(\alpha, \beta, \xi, (1 - \Delta)f, Q, 0) + (1 - 2q)J(\alpha, \beta, \xi, f, Q, 0) + qJ(\alpha, \beta, \xi, (1 + \Delta)f, Q, 0)}.
\]

It follows from (54) 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 which agent he actually observes. Moreover, from (62) and it follows that an agent observing poor, average, and wealthy agents consuming based on posterior beliefs \( \hat{p} \) will have posterior belief

\[
\hat{p} = R(\alpha, \beta, \xi, (1 + \Delta)f),
\]

where \( \beta \) equals \( S_\tau((1 - \Delta)f \tilde{p}), S_\tau(f \tilde{p}), \) and \( S_\tau((1 + \Delta)f \tilde{p}) \) with probabilities \( q, 1 - 2q \), and \( q \), respectively. The fixed point problem that matches aggregate posterior beliefs with agents’ updating is therefore (53).
Existence of a solution to (53) follows from the easily verifiable fact that $R(p, S_r(g \times 0), \xi, f) > 0$ and $R(p, S_r(g \times 1), \xi, f) < 1$, regardless of $g, p, f \in (0, 1)$, and $\xi > 0$. Therefore, the r.h.s of (53), 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.1.

It is straightforward to verify that the function $R$ satisfies $\frac{\partial R}{\partial z} > 0$, since $\frac{\partial R}{\partial z} = \frac{\xi (1 + c)}{2f(1 + \xi)}$, where

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

implying positivity of the derivative. It also follows that $\frac{\partial^2 R}{\partial z^2} = -\frac{2(1-f)(1-p)\xi^2}{((1+f^2+\xi z)^2-4f(1+\xi)(p+\xi))^\frac{3}{2}} < 0$, 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 $\kappa_1 = 0 \iff f = 0$ on $f \in [0, 1]$, for the smooth function $\kappa_1$. Thus, $R$ is a monotone function for positive $0 \leq f \leq 1$. A Taylor expansion of $\kappa_1$ in $f$ around $f = 0$ implies that $\kappa_1$ is on the form $R' = -c_1 f^2 + O(f^3)$, where the constant $c_1 > 0$, altogether implying that $\frac{\partial R}{\partial f} < 0$ for small positive $f$, and thereby for all $0 \leq f \leq 1$ (since $R$ is monotone).

Now, we use these properties of $R$ to show that the total derivative of the r.h.s. of (53) w.r.t. $\Delta$ is negative. Specifically, using the notation $R_i$ for the partial derivative of the function $R$ w.r.t. its $i$th argument, from the calculus of total derivatives it follows that this r.h.s. derivative is on the form

$$q\bar{p}f(-S_r(f\bar{p}(1 - \Delta)) R_2(\cdot) + S_r(f\bar{p}(1 + \Delta)) R_2(\cdot)) + f(q R_4(\cdot) + (1 - 2q) R_4(\cdot) + q R_4(\cdot)).$$

Now, since $R_4(\cdot) < 0$, the second part of this expression is negative. Moreover, $S_r$ 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 (53) is decreasing in $\Delta$.

Because $R$ is increasing and concave in its second argument, it follows that the r.h.s. of (53) is concave and increasing in $\bar{p}$, and since $R(p, 0, \xi, f) > 0$, 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 (53).

Proof of Proposition 11: The proof of the Bayesian updating follows similar lines as in
Proposition C.2. Define
\[
J(\alpha, \beta, \xi, f, g, Q, x) = \int_0^1 t^{\alpha Q - 1 + x} (1 - t)^{(1-\alpha)Q-1} (g + ft)^{\beta \xi Q} (1 - g - ft)^{(1-\beta)\xi Q} dt. \tag{55}
\]

Standard properties of Beta distributions, implies that an agent’s posterior estimate is
\[
\hat{p} = R(\alpha, \beta, \xi, f, g) \overset{\text{def}}{=} \lim_{Q \to \infty} \frac{J(\alpha, \beta, \xi, f, g, Q, 1)}{J(\alpha, \beta, \xi, f, g, Q, 0)}, \tag{56}
\]
when \( Q \) is very large. Moreover, taking the derivative of the term inside the integral of (55) 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:
\[
\frac{\alpha}{\hat{p}} - \frac{1 - \alpha}{1 - \hat{p}} + f \left( \frac{\beta \xi}{g + f \hat{p}} - \frac{1 - \beta \xi}{1 - g - f \hat{p}} \right) = 0 \tag{57}
\]
for large \( Q \). Substituting in the equilibrium condition \( \hat{p} = \bar{p}, \bar{p} \) as a function of \( r \) defined in (38), setting \( \alpha = p, \beta = S_r (g + f \bar{p}), \xi = 1, f \) and \( g \) as defined in (33), and solving for \( \tau \) in (57) leads to the functional relation:
\[
\tau(p, r) = \left( 2p + 2p^2 - 32r + 20pr + 4p^2 r + 36r^2 + 3pr^2 + \right.
\]
\[
+ 5p^2 r^2 - 50r^3 - 5pr^3 + 3p^2 r^3 + 25r^4 - 10pr^4 + p^2 r^4 \right)
\]
\[
/ \left( -2p + 2p^2 + 20pr + 4p^2 r + 48r^2 + 7pr^2 + \right.
\]
\[
+ 5p^2 r^2 - 54r^3 - pr^3 + 3p^2 r^3 + 15r^4 - 8pr^4 + p^2 r^4 \right). \tag{58}
\]

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 \( \tau \) is strictly increasing in \( r \) in a neighborhood of the origin, regardless of \( p \), and that \( \tau \) approaches infinity for some \( r(p) < 1/2 \), so equilibrium is defined for all parameter values \( p \) and \( \tau \), and \( r \) increases in \( \tau \), as does then \( \bar{p} \). Finally, since \( \bar{c}_0^l \) is decreasing in \( r \), see (35), and \( \bar{c}_0 = \bar{c}^e - \bar{c}_0^l \), it follows that \( \bar{c}_0 \) is increasing in \( r \), and then also in \( \tau \), since \( r \) is increasing in \( \tau \).

\[ \blacksquare \]
B Other Utility Functions

The combination in the base model of the utility specification in (1), which leads to a linear consumption function in wealth (5), and the assumption that when \( W/2 \) is consumed at time 0 all bins contain consumption, makes the relationship between \( \bar{p} \) and \( E[z] \) in (10) especially tractable, which allows for a strong characterization of equilibrium.

We now verify that qualitatively similar results as in Proposition 1 also hold under more common utility specifications. For example, consider the case in which agents have power utility,

\[
U = c^{\frac{1}{1-\gamma}} + \frac{1}{(1-\gamma)} c, \quad \text{with \( \gamma \geq 1 \) (where in the case \( \gamma = 1 \), log-utility is used)}.
\]

The consumption shock, \( \epsilon \) is assumed to take on value \( W/2 \) 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

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

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 case of of power utility, \( G \) is a nonlinear function for which a closed form solution is not available, bare a few special values of \( \gamma \).\(^{\text{35}}\) However, the following behavior of \( G \) is easy to show:

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

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

**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. \( \blacksquare \)

If an agent observes a fraction \( x \) of consumption bins, his posterior expected value of \( p \) is

\[
\hat{p} = \frac{p + \xi G^{-1}(x)}{1 + \xi}.
\]

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

\(^{\text{35}}\)For the special case when \( \gamma = 1 \), the closed form solution is \( G(\bar{p}) = \frac{5}{4} \left( 5 - \bar{p} - \sqrt{9 - 10 \bar{p} + \bar{p}^2} \right) \).
The fixed point equilibrium condition:
\[
\bar{p} = \frac{p + \xi G^{-1}(S_{\tau}(G(\bar{p})))}{1 + \xi}.
\] (59)

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

**Proposition B.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
\[
G^{-1}(y) = \frac{p + \xi G^{-1}(S_{\tau}(y))}{1 + \xi}.
\] (60)

To show the existence of an \( y \in (G(p), 1) \) solving (60), we note that
\[
p = G^{-1}(G(p)) < \frac{p + \xi G^{-1}(G(p))}{1 + \xi} < \frac{p + \xi G^{-1}(S_{\tau}(G(p)))}{1 + \xi},
\]
and that
\[
1 = G^{-1}(G(1)) > \frac{p + \xi G^{-1}(S_{\tau}(G(1)))}{1 + \xi} = \frac{p + \xi}{1 + \xi}.
\]

By the intermediate value theorem, there is therefore a \( y \in (G(p), 1) \) that solves (60), with associated equilibrium probability estimate \( \bar{p} = G^{-1}(y) > G^{-1}(G(p)) = p \).

**C The model with different fraction 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 in subsequent sections to analyze situations where there is heterogeneity in consumption behavior, for example, because of wealth dispersion. Specifically, if a rich agent with probability estimate \( \bar{p} = 1 \) 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 (9).

The following proposition covers the case when \( f < 1 \):
Proposition C.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

$$
\hat{p} = R(p, z, \xi, f) = \frac{1}{2f(1+\xi)} \left( 1 + fp + f\xi + z\xi - \sqrt{(1 + fp + f\xi + z\xi)^2 - 4f(1+\xi)(p + z\xi)} \right).
$$

Proof: Define

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

Using standard properties of Beta distributions, it follows that

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

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

It is easy to verify that when $f = 1$, (61) reduces to the base model formula, $\hat{p} = \frac{p + \xi z}{1 + \xi}$. 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:

$$
\bar{p} = R(p, S_\tau(\bar{p}f), \xi, f).
$$

We now have

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

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

where the function $B$ is defined in (12,13).

Proof: Substituting in the definition of $R$ into the fixed point problem (63) 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 $\tau$. Both lead to more overconsumption in equilibrium.

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


References


Enke, B. (2017, September). What you see is all there is. Working paper, Harvard University.


