Libertarian Paternalism, Information Production, and Financial Decision-Making*

Bruce Ian Carlin†
Simon Gervais‡
Gustavo Manso§

June 6, 2012

Abstract

We develop a theoretical model to analyze the effects of libertarian paternalism on information production and financial decision-making. Individuals in our model appreciate the information content of the recommendations made by a social planner. This affects their incentive to gather information, and in turn the speed at which information spreads across market participants, via social learning or formal advice channels. We characterize situations in which libertarian paternalism improves welfare and contrast them with scenarios in which this policy is suboptimal because of its negative impact on the production and propagation of information.

*We would like to thank Anat Admati, Brad Barber, Tony Bernardo, Zvi Bodie, Patrick Bolton, Michael Brennan, James Choi, Mike Fishman, Willie Fuchs, Larry Glosten, Brett Green, David Hirshleifer, Louis Kaplow, David Laibson, Alessandro Lizzeri, Brigitte Madrian, Randall Morck, Terry Odean, Marcus Opp, Paul Pfleiderer, Adriano Rampini, Francesco Sangiorgi, Tano Santos, Steve Tadelis, Hong Yan, Jeff Zwiebel, and seminar participants at Duke University, the University of South Carolina, the University of Colorado at Boulder, Yale University, Northwestern University, Imperial College, University College London, Princeton University, University of Illinois at Urbana-Champaign, the University of British Columbia, DePaul University, Brigham Young University, Oxford University, Nottingham University, Aalto University, the University of California at Berkeley, Stanford University, Columbia University, Harvard Law School, the Federal Reserve Board, the 2009 conference of the Financial Research Association, the 2010 USC-UCLA Finance Day, the 2011 meeting of the American Economic Association, the 2011 meeting of the Financial Intermediation Research Society, and the 2011 meeting of the Western Finance Association for their comments and suggestions. All remaining errors are the authors' responsibility.

†Anderson School of Management, University of California, Los Angeles, 110 Westwood Plaza Suite C413, Los Angeles, CA 90995, bruce.carlin@anderson.ucla.edu, (310) 825-7246.

‡Fuqua School of Business, Duke University, One Towerview Drive, Durham, NC 27708-0120, sgervais@duke.edu, (919) 660-7683.

§Haas School of Business, 545 Student Services #1900, University of California at Berkeley, Berkeley, CA 94720-1900, manso@haas.berkeley.edu, (510) 643-6623.

1 Introduction

Financial sophistication has lagged behind the growing complexity of retail markets (e.g., NASD Literacy Survey, 2003), degrading personal welfare in the process.[1] What to do about this disparity between complexity and sophistication has received much attention, but the optimal solution remains hotly debated. Whereas some are proponents of increasing awareness through education (e.g., Lusardi and Mitchell, 2007), others favor improving peoples' choices through default options that automatically implement a well thought-out course of action for a subset of individuals. Indeed, libertarian paternalism, as posed by Thaler and Sunstein (2003, 2008), makes sense in many venues and has been shown to improve some of the financial decisions that people make (Thaler and Benartzi, 2004).

Libertarian paternalism is provocative because it links two ideas that are on the surface contradictory, but may amount to an uncompromising compromise. The social planner directs individuals through default options that they are free to use or ignore, so that everyone may enjoy the best of both worlds: guidance without the tax of obtrusion. As such, libertarian paternalism is a convenient hybrid solution between government intervention and free markets, whereby centralized and decentralized uses of information can coexist to maximize welfare.

However, just like market socialism neglects the negative impact of government intervention on the production of knowledge (Hirshleifer, 1973; Stiglitz, 1994), libertarian paternalism may also adversely affect the production and exchange of information that is relevant for financial decision-making. That is, if the amount of information were given exogenously, then libertarian paternalism should reach an optimal balance between centralized and free-market uses of that information. However, if libertarian paternalism reduces information acquisition incentives and in turn the pace of social learning, then it may in fact decrease welfare.

In this paper, we develop a theoretical model to analyze this tradeoff or, more specifically, the

[1] See Carlin (2009) and Carlin and Manso (2011) for further discussion.

net effect of default options on total welfare. Our analysis is grounded in the idea that default options provide information to market participants. As documented empirically by Madrian and Shea (2001), this information reduces the willingness of individuals to educate themselves about the choices available to them and ultimately changes the financial decisions they make. In fact, individuals seldom question the suitability of default options and frequently interpret them as the recommended course of action (Brown and Krishna, 2004; McKenzie et al., 2006). This in turn may reduce the usefulness of the information that people share through social interactions (Duflo and Saez, 2002, 2003; Sorensen, 2006; and Beshears et al., 2010), which have been shown to affect a variety of decisions: choices to participate in markets (Hong et al., 2004; Brown et al., 2008; Kaustia and Knüpfer, 2012), to enroll in retirement plans (Madrian and Shea, 2001; Beshears et al., 2010), and to buy stocks (Shiller and Pound, 1989).[2]

Our analysis considers both a setting in which information percolates according to a social learning technology (e.g., Ellison and Fudenberg, 1993 and 1995; Manski, 2004; Duffie and Manso, 2007), and a setting in which uninformed individuals can purchase information from informed ones (i.e., an advice market). We characterize the scenarios in which libertarian paternalism improves welfare and contrast them to scenarios in which it is suboptimal as a result of its negative impact on the production and aggregation of information. Our work leads to the conclusion that libertarian paternalism should be used judiciously rather than as a blanket policy. Indeed, as implied by Arrow's (1994) arguments about social knowledge, it is important to weigh the social multiplier effects of learning (e.g., Glaeser et al., 2003) when considering the design of default options or more generally the adoption of policies based on libertarian paternalism.[3]

In our base model, each individual must make a financial decision whose payoff depends on his unknown type. The social planner has access to a noisy signal about the average type of

[2] For a survey of the literature on social interactions, see Manski (2000).

[3] Similarly, Ahdieh (2011) stresses the importance for any public intervention aimed at individuals to internalize the social dynamics that it may affect. Also, Camerer et al. (2003) propose a form of asymmetric paternalism aimed at minimizing the externality distortions of regulation.

individuals in the economy. Before receiving this signal, she must decide between two policies: (i) institute a default option that will implicitly disclose her information; (ii) let individuals make their own choices without guidance from an informative default. Individuals who are attentive can exert costly effort to find out their own type, so that they make an even better decision. Higher aggregate effort also increases the probability that any one of these individuals becomes informed through social interactions. Inattentive individuals, on the other hand, are either placed in the default, if it exists, or are stuck with alternative (and possibly biased) choices when left to their own devices.[4]

We derive conditions under which default options are optimal and describe when they destroy social surplus. The tradeoff revolves around the fact that the information contained in the default option provided by the social planner reduces each individual's incentive to gather and share any additional information. Thus, although the information in the default is useful to any one individual, it reduces the positive externalities associated with social learning. When the information-sharing technology is sufficiently effective, the cost of information acquisition is low, and/or the individual-specific information is more valuable, providing a default option is suboptimal. Under these conditions, a social planner maximizes welfare by letting market participants fend for themselves and allowing social learning to thrive. Further, when the planner's information is imprecise, the fraction of attentive individuals is high, or the choices that inattentive people make without guidance are nearly rational, a default is also more likely to decrease social welfare.

These results shed light on when libertarian paternalism is likely to add value. For example, default options are likely to be welfare-improving when individuals are sufficiently homogeneous. Consider the default option of a low-fee life cycle fund that automatically reallocates wealth to fixed income assets as investors age. It is unlikely that there is much variation in preferences for such age-dependent reallocations. Yet, people's ability to access this information for themselves

[4]As we discuss in the paper, such inattentiveness may arise rationally because of a competing concern, or behaviorally due to procrastination, an omission bias, or a status quo bias. We include such agents in the economy to make the model realistic, but their presence actually weakens our results about the negative impact of default options.

is limited. Therefore, in this case, providing a default option is likely to add value. However, default options are unlikely to increase social welfare when people's needs are more heterogeneous or when the information acquired by individuals is relatively valuable compared to the information contained in the default option. An example of this might be a decision about the purchase of a life annuity. People's needs for these retirement vehicles are quite variable (e.g., simple life versus joint survivorship) and given the degree of adverse selection associated with such choices, these decisions are difficult to reverse ex post. Getting the choice right on the first attempt is valuable: if providing defaults for this decision decreases some people's incentives to become savvy, this may lead to a drop in welfare.

We then characterize an economy in which information sales (i.e., advice) are allowed to take place. A fraction of the individuals are recognized as information gatherers, whereas the remainder rely on advice markets for guidance. The social planner faces the same problem as before, and information gatherers decide how much costly effort to employ in accumulating knowledge. The social learning technology then takes the form of information exchanges between information gatherers and the rest of the public.[5] In this version of the model, the presence of a default option decreases the value of advice. That is, since fewer information gatherers become knowledgeable, the quality of advice in the market suffers. As in the base model, not offering a default option sometimes dominates issuing a default option, especially if the cost of effort is low and the value of individual-specific (social planner) information is high (low).

Our work adds to a growing literature that suggests that libertarian paternalism comes with more tradeoffs than initially anticipated. For example, Glaeser (2006) suggests that in some contexts libertarian paternalism may be hard to publicly monitor and may lead to hard paternalism.[6] He also warns about the possibility that social planners are not immune from making errors or having

[5]Because individuals learn to make better decisions by interacting with their skilled peers, our approach is similar in spirit to work by Glaeser (1999) and Glaeser and Maré (2001) in which agents become more productive when working with others who are skilled.

[6]See also Rostbøll (2005), and Whitman and Rizzo (2007) for similar arguments.

biases, which may affect the value of default options. Mitchell (2005) questions the redistributive consequences of libertarian paternalism. Korobkin (2009) argues that, because libertarian paternalism ignores the externalities that individuals create for each other, its policies may not maximize collective welfare even though they induce individuals to make optimal decisions for themselves. Finally, Baker and Lytton (2010) question the logic of leaving the decision to opt out of defaults in the hands of the same biased individuals whom defaults are meant to protect.

Our work also adds to the work spawned by the original debate about socialism between Taylor (1929), Lange (1938), Lerner (1938), and Hayek (1945). In his critique of market socialism, Hayek (1945) argues that more efficient economic allocations result from a decentralized use of knowledge. The subsequent failure of socialistic economies and development of formal theories of information economics eventually led Stiglitz (1994, 2000) to conclude that socialistic approaches to governing are essentially a thing of the past. Our paper illustrates that the incentive effects of public information disclosures (e.g., Hirshleifer, 1971; Burguet and Vives, 2000; Angeletos and Pavan, 2007), the role of markets in producing information (e.g., Grossman, 1976; Grossman and Stiglitz, 1980), and the government's ability or inability to coordinate economic activity (e.g., Bolton and Farrell, 1990; Morris and Shin, 2002; Alonso, Dessein and Matouschek, 2008) all affect the viability of default options and libertarian paternalism.

To our knowledge, the only existing theoretical models of default options are those of Choi et al. (2003) and Carroll et al. (2009). These models analyze the optimal implementation of default options when individuals have a tendency to procrastinate in the face of important economic decisions. A sensible default then serves to mitigate the downfall of procrastinating individuals and, when the tendency to procrastinate becomes excessive, an appropriately chosen bias in the default (i.e., an "offset default") partially restores incentives to act quickly. By emphasizing the informational aspect of default options, our paper complements rather than contradicts these papers' results. Indeed, the inattentive individuals in our model can be thought of as the procrastinators in theirs.

The remainder of the paper is organized as follows. Section 2 outlines our base model of social learning and determines when it is optimal to use default options. In Section 3, we turn to information sales to endogenize the mechanism underlying information propagation. Finally, Section 4 provides some concluding remarks. All proofs are in Appendix A.

2 Libertarian Paternalism with Social Learning

2.1 The Model's Setup

The economy is composed of a social planner and a continuum (a non-atomic finite measure space $(I, \mathcal{I}, \gamma)$) of heterogeneous individuals who all face a significant financial decision. Examples of such a decision might be an investment-consumption choice, a capital allocation decision, or a choice of insurance. For simplicity, but without loss of generality, we set the total measure $\gamma(I)$ of individuals to 1 (i.e., a unit mass).

The ex post utility from the decision for each individual $i \in I$ is given by

$$\tilde{U}_i(x_i) = -(\tilde{\tau}_i - x_i)^2, \tag{1}$$

where $x_i \in \mathbb{R}$ is a choice variable and $\tilde{\tau}_i$ is the individual's true (but unknown) type. That is, the best possible decision that individual i can make is $x_i = \tilde{\tau}_i$, but only individuals who learn their own type can make such a decision. Otherwise, as (1) is a quadratic loss function, the goal of each individual is to choose x_i to be as close to $\tilde{\tau}_i$ as possible in order to minimize his expected loss.

Individuals share a common mean type of $\tilde{\mu}$ that is normally distributed with mean zero and variance Σ_μ. For example, $\tilde{\mu}$ could represent the average optimal savings rate of a given population. Conditional on $\tilde{\mu}$, the type $\tilde{\tau}_i$ of an individual i is normally distributed with mean $\tilde{\mu}$ and variance Σ. To capture the possibility that the optimal decision of an individual is related to that of other individuals in the population, we also assume that $\text{Cov}(\tilde{\tau}_i, \tilde{\tau}_j \mid \tilde{\mu}) = \rho\Sigma$, with $\rho \in [0, 1)$, for any $\{i, j\} \in I^2$ with $i \neq j$.

Individuals vary with regard to how attentive they are when making choices. Specifically, they are *attentive* with probability $\lambda \in (0,1]$, and *inattentive* otherwise. Attentive individuals rationally use all available means to learn about their own type and optimize all of their choices. Inattentive individuals, on the other hand, completely avoid the problems they face and make systematic mistakes unless they are offered some guidance. Lack of attentiveness may arise from the presence of other serious choices that compete for a rational individual's attention, from a simple lack of skills, or from behavioral tendencies (e.g., procrastination, omission bias, or status quo bias). As we will see shortly, the presence of such inattentive individuals is not necessary to derive our results. While adding them makes the model more realistic, we will show that our results are strongest when $\lambda = 1$.

Before choosing x_i, each attentive individual i can exert some effort to learn about his own type. An individual's effort of $e_i \in [0,1]$ comes with a personal utility cost of

$$C(e_i) = \frac{c}{2} e_i^2, \tag{2}$$

where c is a positive constant. Going forward, we assume that $c > 2(\Sigma + \Sigma_\mu)$, which guarantees an interior solution to the effort problem but does not affect the economics of the analysis.[7]

An individual who selects an effort level e_i observes his true type $\tilde{\tau}_i$ with probability

$$e_i + \alpha \bar{e}, \tag{3}$$

where $\bar{e} \equiv \int_I e_i d\gamma$ and $\alpha \in [0,1)$, and observes nothing otherwise. Individuals know when they did not receive an informative signal. Given that $\bar{e}$ represents the average effort exerted by individuals in the population, the signal specification in (3) implies that an individual is more likely to learn his own type when many individuals seek to learn theirs. This positive externality of effort captures the idea that as more people exert effort and more of the population becomes informed, their interactions lead to more spillovers in the learning process. This ultimately makes it easier for attentive agents to learn about the decision that they have to make. As such, the parameter α

[7]Note that a more general cost function $C(e)$ that is increasing and convex, and that satisfies $C(0) = 0$, $C'(0) = 0$, and $\lim_{e \to \frac{1}{2}} C'(e) = \infty$, would lead to the same results, but would greatly hinder tractability.

measures the degree of this information externality.

While we use this reduced-form model for parsimony, it accommodates simple microfoundations. For example, after exerting effort e_i, each attentive individual i has a probability α of meeting some other individual randomly drawn from the population. The meeting of two attentive individuals allows them to jointly produce more information about their problem than the combined amount of information they produce on their own. In particular, an attentive individual i who meets another attentive individual j finds out about his type $\tilde{\tau}_i$ with probability $e_i + e_j$. On the other hand, an attentive individual i who does not meet another attentive individual finds out about his type $\tilde{\tau}_i$ with probability e_i only. Ex ante, then, each attentive individual observes his type with the probability in (3).

Inattentive individuals do not conduct any information acquisition, nor do they make any active decision. Without any intervention by the social planner, each inattentive individual i ends up with the same $x_i = \hat{x}_{\mathrm{N}} \in \mathbb{R}$. Because $\mathrm{E}[\tilde{\tau}_i] = 0$, an individual's inattentiveness is more costly as $\hat{x}_{\mathrm{N}}$ departs from zero.[8] Such departures might result from behavioral heuristics that an individual has learned previously, or may be implied by the prevailing economic or legal situation. For example, an individual who ignores all retirement planning decisions will most likely end up saving less than needed.

The social planner can affect the decisions and outcomes of individuals by instituting a default decision $\hat{x}_{\mathrm{D}}$ that they are free to modify. That is, when a default option is provided, an individual i ends up with $x_i = \hat{x}_{\mathrm{D}}$ unless he proactively chooses a different x_i. As in the work of Thaler and Sunstein (2003, 2008), such "nudges" serve to reduce the incidence and importance of the mistakes that inattentive individuals make. In our model, because these individuals do not pay any attention to their particular financial decisions, they end up adopting the defaults without questioning their merit.

[8] Our results are unaffected if we assume that every inattentive individual picks x_i randomly. In this alternative specification, the mean and variance of this random choice then jointly determine how damaging inattentiveness is.

For this *default option* to be useful, however, it must incorporate some pertinent information about the optimal decision that individuals should make. For this purpose, we assume that the social planner costlessly observes a noisy signal $\tilde{s} = \tilde{\mu} + \tilde{\epsilon}$, where $\tilde{\epsilon}$ is normally distributed with mean zero and variance Σ_ϵ, and is independent from $\tilde{\mu}$ and $\tilde{\tau}_i$ for all $i \in I$. For example, this could correspond to the planner having an informed opinion about the optimal average savings rate for a group of individuals.

The planner is not obligated to help. Instead, she chooses whether to set a default option that takes $\tilde{s}$ into account or to leave individuals to their own devices. The planner's goal in this choice is to maximize total welfare. An important aspect of this decision is the information that the default option conveys to attentive individuals, as empirically documented by Madrian and Shea (2001). Since these individuals are fully rational, they are able to glean information about $\tilde{s}$ from a default option if it is offered. This in turn affects their choice of effort in gathering further information. As we show next, this can have important welfare repercussions.

2.2 Equilibrium and Welfare Analysis

We start our analysis by solving for the social planner's optimal choice of a default $\hat{x}_{\mathrm{D}}$ when she elects to make one available. Her choice takes her information into account, and so reveals $\tilde{s}$ to attentive individuals who are then free to change their own x_i. As such, the benevolent planner's choice of a default simply requires her to maximize the welfare of inattentive individuals who will stick to this default.

Lemma 1. *When offering a default, the central planner chooses* $\hat{x}_{\mathrm{D}} = \delta\tilde{s}$*, where* $\delta \equiv \frac{\Sigma_\mu}{\Sigma_\mu + \Sigma_\epsilon}$.

Let $\mathcal{S}_i^x$ denote the information set of an attentive individual i at the time he must make his decision x_i. This set is equal to $\{\tilde{\tau}_i\}$ if the individual observes his true type, whether or not the social planner sets a default option.[9] When there is a default option and the individual does not

[9]Technically speaking, the information set is $\{\tilde{s}, \tilde{\tau}_i\}$ when the social planner announces a default option and individual i observes his own type, but the additional information provided by $\tilde{s}$ (i.e., knowing $\tilde{s}$ and $\tilde{\tau}_i$ separately) is not useful for any of the decisions that this individual must make.

observe his type, $\mathcal{S}_i^x = \{\tilde{s}\}$. Finally, when there is no default option and the individual does not observe his type, $\mathcal{S}_i^x = \varnothing$. The following lemma defines the optimal choice of x_i, given the information set $\mathcal{S}_i^x$.

Lemma 2. *The optimal choice of x_i for attentive individual i is* $\mathrm{E}\big[\tilde{\tau}_i \mid \mathcal{S}_i^x\big]$.

Before choosing x_i but after the social planner's decision to announce a default option, each attentive individual i chooses the effort level e_i that maximizes his expected utility. This choice takes into account the fact that he will subsequently choose x_i according to Lemma 2. It also depends on individual i's information set $\mathcal{S}_i^e$ at that time, which is then $\{\tilde{s}\}$ if the planner makes a default option available and is empty otherwise. The following lemma summarizes and simplifies this maximization problem.

Lemma 3. *If no default is offered, attentive individual i chooses his effort level e_i to maximize*

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i)\Big] = -(1 - e_i - \alpha\bar{e})\big[\Sigma_\mu + \Sigma\big] - \frac{c}{2}e_i^2. \tag{4}$$

If a default is offered, attentive individual i chooses his effort level e_i to maximize

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i) \mid \tilde{s}\Big] = -(1 - e_i - \alpha\bar{e})\Big[(1-\delta)\Sigma_\mu + \Sigma\Big] - \frac{c}{2}e_i^2. \tag{5}$$

This result highlights the tradeoff faced by each individual. Effort is costly (second term in (4) and (5)) but it reduces the variance that the individual is subject to (first term in (4) and (5)). At the same time, the concerted effort of every individual (as measured by $\bar{e}$ which, as we show below, will be different in the two scenarios) creates a public good that takes the form of a further variance reduction. The first term in (4) and (5) also highlights the informational role of the default option. When individual i fails to learn $\tilde{\tau}_i$ (this happens with probability $1 - e_i - \alpha\bar{e}$), the information contained in $\tilde{s}$ allows him to make a better uninformed choice of x_i (it is then optimal to stick with $x_i = \hat{x}_{\mathrm{D}}$, in fact) than without a default. This is why the term in square brackets is smaller in (5) than in (4).

Of course, this smaller residual variance in the presence of a default has an incentive effect on the attentive individuals. The following proposition characterizes their effort choice, with and without a default option.

Proposition 1. *If the social planner does not adopt a default option, each attentive individual chooses effort*

$$e_i = \frac{\Sigma_\mu + \Sigma}{c} \equiv e^{\mathrm{N}}, \tag{6}$$

and the average effort level of the population is $\bar{e} = \lambda e^{\mathrm{N}}$*. An attentive individual* i *who observes a fully informative signal opts out of the default option and chooses* $x_i = \tilde{\tau}_i$*. An attentive individual* i *who does not become informed chooses* $x_i = 0$*. All inattentive individuals choose* $x_i = \hat{x}_{\mathrm{N}}$*.*

If the social planner implements a default option, each attentive individual chooses effort

$$e_i = \frac{(1-\delta)\Sigma_\mu + \Sigma}{c} \equiv e^{\mathrm{D}}, \tag{7}$$

where $\delta = \frac{\Sigma_\mu}{\Sigma_\mu + \Sigma_\epsilon}$*, and the average effort level of the population is* $\bar{e} = \lambda e^{\mathrm{D}}$*. An attentive individual* i *who observes a fully informative signal opts out of the default option and chooses* $x_i = \tilde{\tau}_i$*. All other individuals choose* $x_i = \hat{x}_{\mathrm{D}} = \delta \tilde{s}$*.*

Inspection of (6) and (7) shows that individuals exert more effort with higher Σ, higher Σ_μ, higher Σ_ϵ, and lower c. That is, the more variance about an individual's type that the acquisition of an informative signal ($\tilde{s}_i = \tilde{\tau}_i$) resolves and the lower the cost of acquisition, the more effort each individual is willing to employ. Importantly, it is also the case that

$$e^{\mathrm{N}} = e^{\mathrm{D}} + \frac{\delta \Sigma_\mu}{c}.$$

This implies that people exert more effort without a default option, and that the difference between e^{N} and e^{D} increases as the social planner's information becomes more useful (i.e., as Σ_μ gets larger, and as Σ_ϵ gets smaller), and as information gathering becomes easier (i.e., as c gets smaller).

The social learning externality $\bar{e}$ comes from the average effort of individuals in the economy. Because only the attentive individuals exert the effort derived in Proposition 1 and all such indi-

Information set $\mathcal{S}_i^x$	With default		Without default
$\varnothing$ (bad)	0	$<$	$1-(1+\alpha\lambda)e^{\mathrm{N}}$
$\tilde{s}$ (better)	$1-(1+\alpha\lambda)e^{\mathrm{D}}$	$>$	0
$\tilde{\tau}_i$ (best)	$(1+\alpha\lambda)e^{\mathrm{D}}$	$<$	$(1+\alpha\lambda)e^{\mathrm{N}}$

Table 1. This table shows the frequency of all the possible information sets $\mathcal{S}_i^x$ that attentive individual i will have at the time he makes his financial decision, x_i.

viduals exert the same effort, $\bar{e}$ is equal to λe^{N} without a default and to λe^{D} with a default. It therefore also follows that there are greater opportunities for people to learn from each other when default options are not provided by the social planner. In this sense, whether a default option is welfare improving depends on the strength of the learning externality relative to the value of the information that the social planner has in her possession.

The essence of this tradeoff is captured in Table 1, which shows the frequency of each possible information set $\mathcal{S}_i^x$ for an attentive individual i at the time he makes his choice of x_i. The first two lines of this table show how default options effectively limit the potential downside of individuals in the economy: individuals never have to make completely uninformed decisions when a default guides their choices. As the third line of the table shows however, the drawback of default options comes in the form of a lower frequency of fully informed decisions. This is particularly important when social learning is potent (i.e., when α is large). As seen in Table 1, the fact that e^{D} is smaller than e^{N} also implies that individuals herd into the default when one is available, as documented by Choi et al. (2002), and Johnson and Goldstein (2003). Indeed, not only is it the case that attentive individuals choose the default with a large probability (as $1-(1+\alpha\lambda)e^{\mathrm{D}}$ is large when e^{D} is small), but so do all the inattentive individuals.

Table 1 abstracts from two additional forces that make the availability of a default option advantageous. One is that the overall cost of information production is greater without a default option, as individuals exert greater effort to produce it. Another is the fact that inattentive individuals are not stuck with their biased choices of $\hat{x}_{\mathrm{N}}$ when the default is set to $\hat{x}_{\mathrm{D}}$ by the social planner; that is, as suggested by Thaler and Sunstein (2003), policies based on libertarian paternalism benefit those who cannot fend for themselves. The following proposition takes all these tradeoffs into account to derive and compare the total welfare with and without a default option.

Proposition 2. *The total welfare without a default option is higher than the total welfare with a default option if the cost parameter c is in the following region:*

$$2(\Sigma_\mu + \Sigma) < c < \frac{\lambda\left(\frac{1}{2} + \alpha\lambda\right)\left[(2-\delta)\Sigma_\mu + 2\Sigma\right]}{1 + (1-\lambda)\frac{\hat{x}_{\mathrm{N}}^2}{\delta\Sigma_\mu}}. \tag{8}$$

This region is non-empty if and only if

$$\frac{\delta\Sigma_\mu}{\Sigma_\mu + \Sigma} < 2 - \frac{2}{\lambda\left(\frac{1}{2} + \alpha\lambda\right)}\left[1 + (1-\lambda)\frac{\hat{x}_{\mathrm{N}}^2}{\delta\Sigma_\mu}\right]. \tag{9}$$

According to Proposition 2, welfare without a default option can be higher than welfare with a default option when the cost of information acquisition is sufficiently low (i.e., when c is sufficiently small).[10] This arises because the presence of a default option reduces people's incentives to learn about the problem they face, which in turn slows the pace of information propagation throughout the economy. Specifically, since the right-hand sides of (8) and (9) are increasing in λ and α, it is better for the social planner to leave the production of knowledge to individuals when many of them have the skills to gather information (large λ) and when this information is easy to communicate to others (large α). In other words, the very presence of a default option creates an incentive for the population to herd into it, a damaging effect when people can easily learn a lot from each other.

The breakdown between attentive and inattentive individuals also plays an important role in the tradeoff of Proposition 2. The right-hand sides of (8) and (9) are increasing in λ: as the fraction of attentive individuals rises, the likelihood that a default option destroys welfare increases. This happens for two reasons. First, the production of knowledge originates with the individuals who are attentive to their financial decision and proactively gather information about it. As mentioned

[10] Recall that c is restricted to be above $2(\Sigma_\mu + \Sigma)$ by assumption in order to avoid corner solutions.

earlier, the information-gathering externality $\bar{e}$ is directly proportional to λ in equilibrium, and so a higher λ leads to more social learning. Second, a large λ also means that there are fewer inattentive individuals whose passive behavior makes them worse off when the planner does not intervene.

Keeping λ fixed, it will also be the case that the welfare of inattentive individuals decreases as their unguided choices become more erratic; that is, the inequalities in (8) and (9) are stricter as $\hat{x}_{\mathrm{N}}$ deviates more from $\mathrm{E}[\tilde{\tau}_i] = 0$. Indeed, when λ is small and $\hat{x}_{\mathrm{N}}^2$ is large, default options protect lots of individuals against the important systematic mistakes that they make on their own. In fact, this is often the main argument behind the use of default options in 401(k) plans (e.g., Thaler and Benartzi, 2004). However, if the mistakes that inattentive people make are relatively minor (i.e., $\hat{x}_{\mathrm{N}}$ is close to zero), this protection is less valuable, especially if it slows down the overall production of information elsewhere in the economy. Because $\mathrm{Var}[\tilde{\tau}_i] = \Sigma_\mu + \Sigma$ and $\mathrm{Var}[\tilde{\tau}_i \mid \tilde{s}] = (1-\delta)\Sigma_\mu + \Sigma$, we can think of $\delta\Sigma_\mu$ as the quantity of risk in each individual's problem that the planner's default eliminates. The fact that $\hat{x}_{\mathrm{N}}^2$ gets normalized by $\delta\Sigma_\mu$ in both (8) and (9) means that the mistakes that inattentive individuals make are important only to the extent that they largely exceed the effectiveness of the social planner in curbing risk.

In a similar vein, the left-hand side of (9), $\frac{\delta\Sigma_\mu}{\Sigma_\mu+\Sigma}$, represents the fraction of risk in $\tilde{\tau}_i$ that is eliminated by the planner's default. When this ratio is small (and $\hat{x}_{\mathrm{N}}^2$ is small relative to $\delta\Sigma_\mu$), the welfare benefit from the planner's guidance is more than offset by the welfare lost from the reduced efficiency with which information is produced at the individual level. Contributing to this ratio being small is a large value of Σ. An interpretation of this result is that Σ is a proxy for the amount of heterogeneity in the population: when people's needs or attributes differ a lot, default options are more likely to be suboptimal. Indeed, when the optimal economic choices of individuals are dispersed, it may be preferable to increase their incentives to gather and exchange information about these choices than to provide a default that makes them content and limit the overall production of knowledge in the economy. Another interpretation is that Σ proxies for the value at risk in each individual's decision: when decisions are more important, the social planner should refrain from issuing a default in order to promote learning and information sharing by individuals.

In short, the absence of a default leads to more cross-sectional variance in choices, but such variance is useful if people's needs vary a lot and social learning is powerful enough for them to jointly produce the information that is necessary to reach optimal economic allocations.

3 Information Sales

So far, our model shows that the information content of default options makes their adoption costly and potentially suboptimal when individuals in the economy can help each other learn about their decisions. In this section, we show that the externality need not be of the form specified in Section 2. In particular, we show that allowing a subset of skilled individuals to sell their information to unskilled individuals can generate similar results. That is, the presence of default options reduces the incentive for individuals to gather and resell their information, potentially leading to a decrease in the overall production of knowledge in the economy and to lower welfare.

To establish our results, we adapt the basic model of Section 2 to a context in which some individuals can (and will) seek the advice of other individuals in the economy. More specifically, we now label the two subsets of individuals as skilled (fraction λ) and unskilled (fraction $1-\lambda$). The set of skilled individuals, which we denote by $I_\lambda \in I$ with $\gamma(I_\lambda) = \lambda$, can gather information about their type with the same technology as before, except that we set $\alpha = 0$ in (3) to emphasize the fact that externalities derive purely from information sales. That is, for a cost of $C(e_i) = \frac{c}{2}e_i^2$, individual $i \in I_\lambda$ receives a signal that reveals his type $\tilde{\tau}_i$ with probability e_i.

As in Section 2, the other individuals $j \in I \backslash I_\lambda$ are unskilled in that gathering information about their own type is prohibitively costly. However, instead of being completely inattentive, these unskilled individuals are now allowed to purchase information from one randomly-picked skilled individual and to rationally use this information to make their financial decision x_j. Although everyone's skill is publicly observable, the private information of any one skilled individual is not.

That is, no one can tell if individual i learned $\tilde{\tau}_i$ or not. Thus, for a price p (to be determined shortly), an unskilled individual j can purchase a signal from a skilled individual i, but does not know if he learns $\tilde{\tau}_i$ (which is correlated with his own type $\tilde{\tau}_j$) or noise (which is not) in the process.[11]

For clarity, we assume throughout this section that the social planner's signal is perfect (i.e., $\Sigma_\epsilon = 0$ so that $\tilde{s} = \tilde{\mu}$), and so the default fully reveals $\tilde{\mu}$ when it is made available. The following lemma characterizes the value derived by an unskilled individual who consults a randomly selected skilled individual for information.

Lemma 4. *If the social planner does not adopt a default option, the maximum amount that an unskilled individual is willing to pay for the information sold by a randomly selected skilled individual is*

$$v_{\mathrm{N}} = \frac{(\Sigma_\mu + \rho\Sigma)^2}{\Sigma_\mu + \Sigma}\bar{e}_\lambda^2, \tag{10}$$

where $\bar{e}_\lambda \equiv \frac{1}{\lambda}\int_{I_\lambda} e_i d\gamma$*. If the social planner adopts a default option, the maximum amount that an unskilled individual is willing to pay for the information sold by a randomly selected skilled individual is*

$$v_{\mathrm{D}} = \rho^2\Sigma\bar{e}_\lambda^2. \tag{11}$$

Unskilled individuals are willing to pay more to learn a skilled individual's information when they know that skilled individuals exert a lot of effort to learn their own type, i.e., v_{N} and v_{D} are both increasing in $\bar{e}_\lambda$. This makes sense as a fraction $\bar{e}_\lambda$ of the λ skilled individuals will be informed in equilibrium, while the other $(1-\bar{e}_\lambda)\lambda$ skilled individuals sell useless noise. From (10) and (11), we can also see that unskilled individuals are willing to pay a higher price for a skilled individual's

[11] We assume that skilled individuals who do not learn their own type sell an uninformative signal that is randomly drawn from a normal distribution with a mean of zero and a variance of $\Sigma_\mu + \Sigma$, which makes it impossible for information buyers to tell noise from real information. The skilled individuals have nothing to gain from doing anything else. Note also that our setup is equivalent to one in which a skilled individual i sells advice to an unskilled individual j in the form of an optimal decision x_j that optimally incorporates his information; that is our results are unaffected by who does the updating of $\tilde{\tau}_j$ given i's information set.

information when their type is more highly correlated with that of other individuals (large ρ); that is, they learn more from others when their financial situation is similar.

For further insight into Lemma 4, let us denote the total variance of $\tilde{\tau}_i$ by $\Sigma_\tau \equiv \Sigma_\mu + \Sigma$ and define $\Gamma \equiv \frac{\Sigma_\mu}{\Sigma_\mu + \Sigma}$. Since the social planner's information about $\tilde{\mu}$ is perfect, this last quantity represents the fraction of the total variance of an individual's type that the default eliminates. Using this notation, it is straightforward to verify that $v_{\text{N}} = \left[\Gamma + \rho(1-\Gamma)\right]^2 \Sigma_\tau \bar{e}_\lambda^2$ and $v_{\text{D}} = \rho^2(1-\Gamma)\Sigma_\tau \bar{e}_\lambda^2$. Keeping the social planner's relative ability to curb risk constant (i.e., keeping Γ fixed), unskilled individuals are willing to pay a higher price for a skilled individual's information when their type is highly variable (large Σ_τ).

This last result is consistent with the fact that, keeping Σ_τ fixed, v_{N} is increasing in Γ, as types are more correlated when the common mean $\tilde{\mu}$ accounts for a larger portion of each individual's type. This is also consistent with v_{D} being decreasing in Γ as, when the social planner announces $\tilde{\mu}$, the unknown portion of an individual's type correlates with someone else's type only to the extent that the default option leaves residual uncertainty. In fact, using (10) and (11), it is straightforward to verify that $v_{\text{N}} > v_{\text{D}}$ for a given total variance Σ_τ and aggregate level of effort $\bar{e}_\lambda$. Indeed, because types are more correlated across individuals when $\tilde{\mu}$ is unknown, it is the case that unskilled individuals are willing to pay more to learn a skilled individual's type when there is no default option offered. As we shall see below, this difference between v_{N} and v_{D} is exacerbated by the fact that the equilibrium effort level of skilled individuals is greater in the absence of a default option.

The price that a skilled individual will end up charging for his information will in general depend on how much competition he faces from other information sellers or, alternatively, on how easy it is for unskilled individuals to consult another skilled individual. To capture these possibilities in a tractable manner, we assume that the economic surplus from a transaction between a skilled individual and an unskilled individual is split as a Nash bargaining outcome. More specifically, we assume that a skilled individual charges $p = \theta\nu_\sigma$ for the information he sells to an unskilled individual, where $\theta \in [0, 1]$ and $\sigma = \text{D}$ if a default option is made available (and $\sigma = \text{N}$ otherwise).

When $\theta = 1$ ($\theta = 0$), the skilled (unskilled) individual extracts all the surplus from the transaction.[12] Setting $\theta \in (0, 1)$ allows us to capture any intermediate market power scenario. Our results are unaffected by the size of θ, as money exchanges between individuals cancel out in the total welfare function that the social planner seeks to maximize.[13] We start with the following result, which describes the equilibrium in the absence of a default option.

Proposition 3. *If the social planner does not adopt a default option, then each skilled individual $i \in I_\lambda$ chooses an effort level $e_i = \frac{\Sigma_\mu + \Sigma}{c} = \frac{\Sigma_\tau}{c}$, and chooses $x_i = \tilde{\tau}_i$ or $x_i = 0$, depending on whether or not he observes $\tilde{\tau}_i$. Each unskilled individual $j \in I \setminus I_\lambda$ purchases a signal $\tilde{s}_j$ (which is $\tilde{\tau}_{\tilde{\imath}}$ or noise) from a randomly selected skilled individual $\tilde{\imath} \in I_\lambda$ for a price $p = \theta v_{\mathrm{N}}$, with v_{N} given by (10), and chooses*

$$x_j = \frac{\Sigma_\mu + \rho\Sigma}{\Sigma_\mu + \Sigma}\bar{e}_\lambda \tilde{s}_j = \big[\Gamma + \rho(1 - \Gamma)\big]\bar{e}_\lambda \tilde{s}_j. \tag{12}$$

The skilled individuals' behavior is the same as in Section 2. In particular, their behavior is not affected by the possibility of reselling their information to unskilled individuals. This is due to the fact that unskilled individuals cannot distinguish between skilled individuals who learn their type and skilled individuals who do not. That is, they pay θv_{N} to the one skilled individual they encounter, informed or not. As we see from (12), the extent to which unskilled individuals rely on the information they purchase depends on its correlation with their type, as increases in ρ, Γ and $\bar{e}_\lambda$ all ultimately lead to a higher correlation between $\tilde{s}_j$ and $\tilde{\tau}_j$. The following result is the analogue of Proposition 3 when the social planner makes a default option $\hat{x}_{\mathrm{D}} = \tilde{\mu}$ available.

Proposition 4. *If the social planner adopts a default option, then each skilled individual $i \in I_\lambda$ chooses an effort level $e_i = \frac{\Sigma}{c} = \frac{(1-\Gamma)\Sigma_\tau}{c}$, and chooses $x_i = \tilde{\tau}_i$ or $x_i = \tilde{\mu}$, depending on whether or not he observes $\tilde{\tau}_i$. Each unskilled individual $j \in I \setminus I_\lambda$ purchases a signal $\tilde{s}_j$ (which is $\tilde{\tau}_{\tilde{\imath}}$ or noise)*

[12] Note that when $\theta = 0$, the transaction can be interpreted as a free information exchange between two individuals with different skills. For example, this captures the situation in which a new employee asks an existing employee of the same firm about his choices in the company's 401(k) plan.

[13] Of course, any welfare improvement from adding or removing a default will be Pareto-dominant for an interior range of θ that appropriately splits the surplus between skilled and unskilled.

from a randomly selected skilled individual $\tilde{\imath} \in I_\lambda$ *for a price* $p = \theta v_{\mathrm{D}}$*, with* v_{D} *given by (11), and chooses*

$$x_j = \tilde{\mu} + \rho \bar{e}_\lambda (\tilde{s}_j - \tilde{\mu}). \tag{13}$$

As in Proposition 3: more risk (large Σ, or large Σ_τ keeping Γ fixed) leads to more effort, and more correlation (large ρ and $\bar{e}_\lambda$) leads to heavier reliance on purchased information. When Γ is large, skilled individuals do not gain much from learning their type perfectly, as the default option already reveals a large portion of their type. As such, they work less. Although Γ affects the price of information (as discussed earlier), it does not affect the weight that unskilled individuals put on the information they acquire from skilled individuals. Instead, they use the default option to remove the common mean component $\tilde{\mu}$ included in the signal and place weight on $(\tilde{s}_j - \tilde{\mu})$ only to the extent that it is correlated with $(\tilde{\tau}_j - \tilde{\mu})$.

Finally, note that as in Proposition 1, the skilled individuals exert a higher level of effort in the absence of a default option since the incentive to gather information is stronger when they do not have a default option to fall back on. This in turn causes the quality of their advice to decrease, and further amplifies the previously discussed difference between v_{N} and v_{D}. That is, unskilled individuals do not benefit as much from a skilled individual's information, and are thus inclined to pay less for it.

As in Section 2, to assess the pros and cons of the planner's default option, we compare total welfare with and without this option. In this case, welfare must be aggregated over skilled and unskilled individuals. This is done in the following lemma.

Lemma 5. *The total welfare without a default option is*

$$W_{\mathrm{N}} = -(\Sigma_\mu + \Sigma) + \frac{\lambda}{2c}(\Sigma_\mu + \Sigma)^2 + \frac{1-\lambda}{c^2}(\Sigma_\mu + \Sigma)(\Sigma_\mu + \rho\Sigma)^2. \tag{14}$$

The total welfare with a default option is

$$W_{\mathrm{D}} = -\Sigma + \frac{\lambda}{2c}\Sigma^2 + \frac{1-\lambda}{c^2}\rho^2\Sigma^3. \tag{15}$$

In Section 2, an increase in α enhances overall welfare through the larger information gathering externalities that individuals have on each other. We can now see from (14) and (15) that increases in ρ have a similar effect in the presence of information sales. More precisely, straightforward differentiation of these two expressions with respect to ρ lead to

$$\frac{\partial W_{\mathrm{N}}}{\partial \rho} = \frac{2(1-\lambda)}{c^2}(\Sigma_\mu + \Sigma)(\Sigma_\mu + \rho\Sigma)\Sigma > 0 \tag{16}$$

and

$$\frac{\partial W_{\mathrm{D}}}{\partial \rho} = \frac{2(1-\lambda)}{c^2}\rho\Sigma^3 > 0. \tag{17}$$

That is, a larger correlation across individuals' types leads to more welfare when a formal advice channel, like information sales, is incorporated. We can also see that the increase in welfare accommodated by this advice channel is more important when a sizeable fraction of the population is unskilled (i.e., $1-\lambda$ is large). Finally, it is clear that (16) is greater than (17): the advice channel is more crucial and the role of ρ greater when the social planner refrains from making a default option available, as unskilled individuals can then rely only on the skilled individuals' information for their decisions.

The next proposition is the analogue of Proposition 2 when we allow for information sales.

Proposition 5. *The total welfare W_{N} without a default option is higher than the total welfare W_{D} with a default option if the cost parameter c is sufficiently small (the bound is shown in the proof) and*

$$\frac{\Sigma}{\Sigma_\mu(\Sigma_\mu + \Sigma)}\Big[(2\rho - 1)\Sigma_\mu + (\rho^2 + 2\rho - 1)\Sigma\Big] > \frac{\lambda}{2(1-\lambda)}. \tag{18}$$

As mentioned above, ρ plays an especially important welfare role in information sales when the social planner does not make a default option available. Proposition 5 formalizes this by showing that the availability of a default option is always optimal when $\rho^2 + 2\rho - 1 < 0$ (i.e., when $\rho \lesssim 0.414$), as this always makes the left-hand side of (18) negative.[14] That is, unskilled individuals

[14] When $\rho^2 + 2\rho - 1 < 0$, we also have $2\rho - 1 < \rho^2 + 2\rho - 1 < 0$, and so both terms in the square brackets in (18) are negative.

are better off learning the common component of their type perfectly from the social planner when the information that can be acquired from other individuals is not all that useful. This implies that default options are especially valuable when the needs of an individual are unlikely to be similar to those of his peers, including the ones who can advise him.

Since (18) can be rewritten as

$$(2\rho-1)\frac{1}{\frac{\Sigma_\mu}{\Sigma}+1}+(\rho^2+2\rho-1)\frac{1}{\frac{\Sigma_\mu}{\Sigma}\left(\frac{\Sigma_\mu}{\Sigma}+1\right)}>\frac{\lambda}{2(1-\lambda)},$$

we can also see from Proposition 5 that default options are less valuable when Σ is large and Σ_μ is small, which is similar to our findings in Section 2. The extent to which the social planner can resolve the uncertainty faced by the population is still an important determinant of the usefulness of the default option. Interestingly, however, default options are more valuable when a larger fraction of the population is skilled (large λ), even when ρ is large. This arises because the information externalities that skilled individuals bring to the market through information sales is limited: the small number of unskilled individuals leads to a small number of information sales, and so the effort choices of skilled individuals with and without a default option (as derived in Proposition 4) do not lead to significantly different externalities.[15]

In sum, because the nudges that come with libertarian paternalism contain useful information, they affect the incentives of those individuals who have other means to learn about their financial decisions. When, as suggested by Hayek (1945), individuals can and do organize to maximize their joint production and use of knowledge through social networks or formal advice channels, these nudges can have negative welfare consequences. Ultimately therefore, every application of libertarian paternalism must come with a careful assessment of the implicit information/incentive tradeoff.

[15] Note that this section's assumption that skilled individuals do not learn from each other (i.e., $\alpha = 0$) directly contributes to this result. More generally, a large number λ of skilled individuals leads to better information production when the externalities across the set of skilled individuals are larger than those across skilled and unskilled (and vice versa for a small λ).

4 Concluding Remarks

Libertarian paternalism is an alluring idea because it allows knowledge to be used by a central planner without explicitly preventing concurrent decentralized uses. At first glance, it appears to be a policy that even Lange, Lerner, and Hayek could all agree upon, and it may indeed work well in many settings. However, as we show in this paper, one needs to be cautious when implementing the ideals of such a policy because libertarian paternalism may alter the production of information in the economy. Moreover, it is not necessarily the paternalistic partner in this union that causes problems in the relationship, but the freedom that participants exercise that may lead to welfare-decreasing side-effects. Indeed, as its name suggests, libertarian paternalism preserves the rights of individuals to act in their own best interest, benefit from each other's effort provision, and shirk in their own responsibilities. In the face of non-cooperative incentives, libertarian paternalism may induce or worsen externalities that decrease welfare, even though it does not explicitly force people to act in a prescribed manner.

In the paper, we analyze a theoretical model to characterize one such distortion: information acquisition and social learning. As documented by Madrian and Shea (2001) in the context of 401(k) plan choices, default options have information content, which participants may take into consideration when making key financial decisions. Importantly, this affects their incentives to gather further information, which in turn may alter the success of information aggregation, either through social learning or through formal information exchanges.

We characterize the situations in which libertarian paternalism is more likely to add or reduce value given this externality. We show that default options tend to improve social welfare when acquiring information is costly, information is not easily shared across individuals, and people are more heterogeneous in their attributes or needs. Based on our model, default options will likely decrease welfare when the social planner knows less about its constituents, when people are heterogeneous, and when the value at stake in the decision is large.

Our theory adds an important tradeoff in the optimal implementation of libertarian paternalism through public recommendations and advice. Further study of the externalities induced by libertarian paternalism are the subject of future research, which appears warranted given the potential welfare import of this policy.

Appendix A. Proofs

Proof of Lemma 1

When choosing $\hat{x}_{\mathrm{D}}$, the social planner seeks to maximize

$$\mathrm{E}\big[\tilde{U}_i(\hat{x}_{\mathrm{D}}) \mid \tilde{s}\big] = \mathrm{E}\big[-(\tilde{\tau}_i - \hat{x}_{\mathrm{D}})^2 \mid \tilde{s}\big] = -\mathrm{E}\big[\tilde{\tau}_i^2 \mid \tilde{s}\big] + 2\hat{x}_{\mathrm{D}}\mathrm{E}\big[\tilde{\tau}_i \mid \tilde{s}\big] - \hat{x}_{\mathrm{D}}^2.$$

Straightforward differentiation with respect to $\hat{x}_{\mathrm{D}}$ yields the first-order condition for this problem, $2\mathrm{E}\big[\tilde{\tau}_i \mid \tilde{s}\big] - 2\hat{x}_{\mathrm{D}} = 0$. This in turn yields $\hat{x}_{\mathrm{D}} = \mathrm{E}\big[\tilde{\tau}_i \mid \tilde{s}\big] = \frac{\Sigma_\mu}{\Sigma_\mu + \Sigma_\epsilon}\tilde{s}$, after a simple application of the projection theorem. It is straightforward to verify that the second-order condition is satisfied. ■

Proof of Lemma 2

Individual i must choose x_i in order to maximize

$$\mathrm{E}\big[\tilde{U}_i(x_i) \mid \mathcal{S}_i^x\big] = \mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \mid \mathcal{S}_i^x\big] = -\mathrm{E}\big[\tilde{\tau}_i^2 \mid \mathcal{S}_i^x\big] + 2x_i\mathrm{E}\big[\tilde{\tau}_i \mid \mathcal{S}_i^x\big] - x_i^2.$$

By differentiating this expression with respect to x_i, we obtain the first-order condition for this problem, $2\mathrm{E}\big[\tilde{\tau}_i \mid \mathcal{S}_i^x\big] - 2x_i = 0$, which yields $x_i = \mathrm{E}\big[\tilde{\tau}_i \mid \mathcal{S}_i^x\big]$. It is straightforward to verify that the second-order condition is satisfied. ■

Proof of Lemma 3

First, let us consider the case without a default option. Using Lemma 2 and the fact that $\mathcal{S}_i^e = \varnothing$, individual i's expected utility is given by

$$\begin{aligned}
\mathrm{E}\Big[\tilde{U}_i(x_i) \mid \mathcal{S}_i^e\Big] &= \mathrm{E}\Big[-(\tilde{\tau}_i - x_i)^2\Big] = \mathrm{E}\Big\{\mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \mid \mathcal{S}_i^x\big]\Big\} \\
&= \Pr\big\{\mathcal{S}_i^x = \{\tilde{\tau}_i\}\big\}\mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \mid \tilde{\tau}_i\big] + \Pr\big\{\mathcal{S}_i^x = \varnothing\big\}\mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2\big] \\
&= (e_i + \alpha\bar{e})\mathrm{E}\big[-(\tilde{\tau}_i - \tilde{\tau}_i)^2\big] + (1 - e_i - \alpha\bar{e})\mathrm{E}\big[-(\tilde{\tau}_i - 0)^2\big] \\
&= -(1 - e_i - \alpha\bar{e})(\Sigma_\mu + \Sigma).
\end{aligned}$$

The result obtains after we subtract the cost of effort $C(e_i)$ for individual i, as given in (2).

Now, let us consider the case with a default option. Using the projection theorem for normal variables, it is straightforward to show that $\mathrm{E}\big[\tilde{\tau}_i \mid \tilde{s}\big] = \frac{\Sigma_\mu}{\Sigma_\mu+\Sigma_\epsilon}\tilde{s} = \delta\tilde{s}$ and $\mathrm{Var}\big[\tilde{\tau}_i \mid \tilde{s}\big] = \Big(1-\frac{\Sigma_\mu}{\Sigma_\mu+\Sigma_\epsilon}\Big)\Sigma_\mu + \Sigma = (1-\delta)\Sigma_\mu + \Sigma$, where $\delta = \frac{\Sigma_\mu}{\Sigma_\mu+\Sigma_\epsilon}$. Thus, when individual i's information set is $\mathcal{S}_i^x = \{\tilde{s}\}$ at the time of his decision about x_i, Lemma 2 implies that $x_i = \delta\tilde{s}$. When individual i observes his type and $\mathcal{S}_i^x = \{\tilde{\tau}_i\}$, then he chooses $x_i = \tilde{\tau}_i$, as before. At the time of his effort decision, individual i's information set is $\mathcal{S}_i^e = \{\tilde{s}\}$, and thus

$$\begin{aligned}
\mathrm{E}\Big[\tilde{U}_i(x_i) \mid \mathcal{S}_i^e\Big] &= \mathrm{E}\Big[-(\tilde{\tau}_i - x_i)^2 \mid \tilde{s}\Big] = \mathrm{E}\Big\{\mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \mid \mathcal{S}_i^x\big] \mid \tilde{s}\Big\} \\
&= \Pr\big\{\mathcal{S}_i^x = \{\tilde{\tau}_i\} \mid \tilde{s}\big\}\mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \mid \tilde{\tau}_i\big] + \Pr\big\{\mathcal{S}_i^x = \{\tilde{s}\} \mid \tilde{s}\big\}\mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \mid \tilde{s}\big] \\
&= (e_i + \alpha\bar{e})\mathrm{E}\big[-(\tilde{\tau}_i - \tilde{\tau}_i)^2\big] + (1 - e_i - \alpha\bar{e})\mathrm{E}\big[-(\tilde{\tau}_i - \delta\tilde{s})^2 \mid \tilde{s}\big] \\
&= -(1 - e_i - \alpha\bar{e})\mathrm{Var}\big[\tilde{\tau}_i \mid \tilde{s}\big] = -(1 - e_i - \alpha\bar{e})\Big[(1-\delta)\Sigma_\mu + \Sigma\Big].
\end{aligned}$$

Therefore, each individual i chooses e_i to maximize

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i) \mid \tilde{s}\Big] = -(1 - e_i - \alpha\bar{e})\Big[(1-\delta)\Sigma_\mu + \Sigma\Big] - \frac{c}{2}e_i^2.$$

This completes the proof. ■

Proof of Proposition 1

The optimal economic decisions of an attentive individual all follow from Lemma 2. In the absence of a default option, Lemma 3 shows that each attentive individual i chooses e_i to maximize

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i)\Big] - (1 - e_i - \alpha\bar{e})\Big[\Sigma_\mu + \Sigma\Big] - \frac{c}{2}e_i^2.$$

The first-order condition for this problem is

$$\Sigma_\mu + \Sigma - ce_i = 0,$$

which implies (6) and $\bar{e} \equiv \int_I e_i d\gamma = \lambda e^{\mathrm{N}} + (1-\lambda)\cdot 0 = \lambda e^{\mathrm{N}}$. It is easy to see that the second order condition is satisfied.

Similarly, if a default is provided, Lemma 3 shows that each attentive individual i chooses e_i to maximize

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i) \mid \tilde{s}\Big] = -(1 - e_i - \alpha\bar{e})\Big[(1-\delta)\Sigma_\mu + \Sigma\Big] - \frac{c}{2}e_i^2.$$

The first-order condition for this problem is

$$(1-\delta)\Sigma_\mu + \Sigma - ce_i = 0,$$

which leads to (7) and to $\bar{e} \equiv \int_I e_i d\gamma = \lambda e^{\mathrm{D}} + (1-\lambda)\cdot 0 = \lambda e^{\mathrm{D}}$. Again, it is easy to verify that the second-order condition is satisfied. ■

Proof of Proposition 2

We can use the effort choices from Proposition 1 in Lemma 3 to compute the welfare of attentive individuals without a default option,

$$W_{\mathrm{N}}^{\mathrm{A}} = -(\Sigma_\mu + \Sigma) + \frac{(1+2\alpha\lambda)}{2c}(\Sigma_\mu + \Sigma)^2, \tag{A1}$$

and with a default option,

$$W_{\mathrm{D}}^{\mathrm{A}} = -\big[(1-\delta)\Sigma_\mu + \Sigma\big] + \frac{(1+2\alpha\lambda)}{2c}\big[(1-\delta)\Sigma_\mu + \Sigma\big]^2. \tag{A2}$$

Inattentive individuals do not change their economic decision from $\hat{x}_{\mathrm{N}}$ or $\hat{x}_{\mathrm{D}}$ (which is set to $\delta\tilde{s}$ by the social planner, according to Lemma 1), and so their welfare can be easily calculated to be

$$W_{\mathrm{N}}^{\mathrm{I}} = -(\Sigma_\mu + \Sigma) - \hat{x}_{\mathrm{N}}^2 \tag{A3}$$

without a default option, and

$$W_{\mathrm{D}}^{\mathrm{I}} = -\big[(1-\delta)\Sigma_\mu + \Sigma\big] \tag{A4}$$

with a default option. Total ex ante social welfare is therefore

$$W_{\mathrm{N}} = \lambda W_{\mathrm{N}}^{\mathrm{A}} + (1-\lambda)W_{\mathrm{N}}^{\mathrm{I}} = -(\Sigma_\mu + \Sigma) + \lambda\frac{(1+2\alpha\lambda)}{2c}(\Sigma_\mu + \Sigma)^2 - (1-\lambda)\hat{x}_{\mathrm{N}}^2 \tag{A5}$$

without a default option, and

$$W_{\mathrm{D}} = \lambda W_{\mathrm{D}}^{\mathrm{A}} + (1-\lambda) W_{\mathrm{D}}^{\mathrm{I}} = -\big[(1-\delta)\Sigma_\mu + \Sigma\big] + \lambda \frac{(1+2\alpha\lambda)}{2c} \big[(1-\delta)\Sigma_\mu + \Sigma\big]^2 \tag{A6}$$

with a default option.

A simple comparison of (A5) and (A6) yields the second inequality in (8). The first inequality in (8) is by assumption. The region is non-empty if and only if

$$2(\Sigma_\mu + \Sigma) < \frac{\lambda\left(\frac{1}{2} + \alpha\lambda\right)\big[(2-\delta)\Sigma_\mu + 2\Sigma\big]}{1 + (1-\lambda)\frac{\hat{x}_{\mathrm{N}}^2}{\delta\Sigma_\mu}},$$

which simplifies to the condition in (9). ■

Proof of Lemma 4

Let $\tilde{s}_j$ denote the information purchased by unskilled individual j from skilled individual i, and let us first consider the case in which the social planner does not make a default option available. After individual j receives $\tilde{s}_j$, we know from Lemma 2 that he chooses

$$x_j = \mathrm{E}\big[\tilde{\tau}_j \mid \tilde{s}_j\big] = \bar{e}_\lambda \mathrm{E}\big[\tilde{\tau}_j \mid \tilde{s}_j = \tilde{\tau}_i\big] + (1-\bar{e}_\lambda)\underbrace{\mathrm{E}\big[\tilde{\tau}_j\big]}_{=0} = \bar{e}_\lambda \beta_j \tilde{s}_j,$$

where $\beta_j \equiv \frac{\Sigma_\mu + \rho\Sigma}{\Sigma_\mu + \Sigma}$ is obtained from the normal projection theorem. Thus, before learning $\tilde{s}_j$ but knowing that purchasing it for a price p will lead to an economic decision x_j, individual j's expected utility is

$$\begin{aligned}\mathrm{E}\Big[\tilde{U}_i(x_j) - p\Big] &= \mathrm{E}\Big[(\tilde{\tau}_j - \bar{e}_\lambda\beta_j\tilde{s}_j)^2\Big] - p \\ &= \bar{e}_\lambda \mathrm{E}\Big[(\tilde{\tau}_j - \bar{e}_\lambda\beta_j\tilde{\tau}_i)^2\Big] + (1-\bar{e}_\lambda)\mathrm{E}\Big[(\tilde{\tau}_j - \bar{e}_\lambda\beta_j\tilde{\eta}_i)^2\Big] - p,\end{aligned} \tag{A7}$$

where $\tilde{\eta}_i$ has the same distribution as $\tilde{\tau}_i$ but is independent from it (and from $\tilde{\tau}_j$). Since

$$\mathrm{E}\Big[(\tilde{\tau}_j - \bar{e}_\lambda\beta_j\tilde{\tau}_i)^2\Big] = (\Sigma_\mu + \Sigma) - 2\bar{e}_\lambda\beta_j(\Sigma_\mu + \rho\Sigma) + \bar{e}_\lambda^2\beta_j^2(\Sigma_\mu + \Sigma)$$

and

$$\mathrm{E}\Big[(\tilde{\tau}_j - \bar{e}_\lambda\beta_j\tilde{\eta}_i)^2\Big] = (\Sigma_\mu + \Sigma) + \bar{e}_\lambda^2\beta_j^2(\Sigma_\mu + \Sigma),$$

we can rewrite (A7) as

$$\mathrm{E}\Big[\tilde{U}_i(x_j) - p\Big] = (\Sigma_\mu + \Sigma) - 2\bar{e}_\lambda^2 \beta_j (\Sigma_\mu + \rho\Sigma) + \bar{e}_\lambda^2 \beta_j^2 (\Sigma_\mu + \Sigma) - p.$$

Finally, after we replace β_j by $\frac{\Sigma_\mu + \rho\Sigma}{\Sigma_\mu + \Sigma}$, this simplifies to

$$\mathrm{E}\Big[\tilde{U}_i(x_j) - p\Big] = \Sigma_\mu + \Sigma - \bar{e}_\lambda^2 \frac{(\Sigma_\mu + \rho\Sigma)^2}{\Sigma_\mu + \Sigma} - p. \tag{A8}$$

If instead individual j decides not to purchase any information, his optimal economic choice is $x_j = 0$ and his expected utility is

$$\mathrm{E}\Big[\tilde{U}_i(x_j)\Big] = \mathrm{E}\Big[\tilde{\tau}_j^2\Big] = \Sigma_\mu + \Sigma. \tag{A9}$$

Thus the largest price p that makes individual j indifferent between purchasing and not purchasing $\tilde{s}_j$ is that which makes (A8) and (A9) equal, as shown in (10). The case in which the social planner makes a default option available is similarly derived. ■

Proof of Proposition 3

Let $\tilde{\pi}_i$ denote the profits that a skilled individual $i \in I_\lambda$ generates from selling information to unskilled individuals. With an information price $p = \theta v_{\mathrm{N}}$, the $1 - \lambda$ unskilled individuals will pay a total sum of $(1-\lambda)p = (1-\lambda)\theta v_{\mathrm{N}}$ to acquire signals from the λ skilled individuals. Since these skilled individuals are randomly selected, the expected profits from information sales of any one skilled individual i are

$$\mathrm{E}\big[\tilde{\pi}_i\big] = \frac{(1-\lambda)\theta v_{\mathrm{N}}}{\lambda}.$$

Thus, using the same notation and reasoning as in Lemma 3, this skilled individual i must choose e_i in order to maximize

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i) + \tilde{\pi}_i\Big] = -(1-e_i)(\Sigma_\mu + \Sigma) - \frac{c}{2}e_i^2 + \frac{(1-\lambda)\theta v_{\mathrm{N}}}{\lambda}.$$

Because the last term in this expression is not affected by this individual's choice of e_i, the first-order and second-order conditions for this maximization problem are identical to those in the proof of Proposition 1, and so lead to $e_i = \frac{\Sigma_\mu+\Sigma}{c}$. After purchasing $\tilde{s}_j$ from a skilled individual, unskilled individual j must choose x_j in order to maximize $\mathrm{E}\big[-(\tilde{\tau}_j - x_j)^2 \mid \tilde{s}_j\big]$. By Lemma 2, this individual chooses

$$x_j = \mathrm{E}\big[\tilde{\tau}_j \mid \tilde{s}_j\big] = \bar{e}_\lambda \mathrm{E}\big[\tilde{\tau}_j \mid \tilde{s}_j = \tilde{\tau}_i\big] + (1-\bar{e}_\lambda)\underbrace{\mathrm{E}\big[\tilde{\tau}_j\big]}_{=0} = \bar{e}_\lambda \frac{\Sigma_\mu + \rho\Sigma}{\Sigma_\mu + \Sigma}\tilde{s}_j,$$

where the last equality is obtained using the projection theorem. Using the fact that $\Sigma_\mu = \Gamma\Sigma_\tau$ and $\Sigma = (1-\Gamma)\Sigma_\tau$, we can rewrite this last expression as $x_j = \big[\Gamma + \rho(1-\Gamma)\big]\bar{e}_\lambda\tilde{s}_j$. ■

Proof of Proposition 4

Let $\tilde{\pi}_i$ denote the profits that a skilled individual $i \in I_\lambda$ generates from selling information to unskilled individuals. With an information price $p = \theta v_{\mathrm{D}}$, the $1-\lambda$ unskilled individuals will pay a total sum of $(1-\lambda)p = (1-\lambda)\theta v_{\mathrm{D}}$ to acquire signals from the λ skilled individuals. Since these skilled individuals are randomly selected, the expected profits from information sales of any one skilled individual i are

$$\mathrm{E}\big[\tilde{\pi}_i\big] = \frac{(1-\lambda)\theta v_{\mathrm{D}}}{\lambda}.$$

Thus, using the same notation and reasoning as in Lemma 3, this skilled individual i must choose e_i in order to maximize

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i) + \tilde{\pi}_i \;\Big|\; \tilde{\mu}\Big] = -(1-e_i)\Sigma - \frac{c}{2}e_i^2 + \frac{(1-\lambda)\theta v_{\mathrm{D}}}{\lambda}.$$

Because the last term in this expression is not affected by this individual's choice of e_i, the first-order and second-order conditions for this maximization problem are identical to those in the proof of Proposition 1, and so lead to $e_i = \frac{\Sigma}{c}$. After purchasing $\tilde{s}_j$ from a skilled individual, unskilled individual j must choose x_j in order to maximize $\mathrm{E}\big[-(\tilde{\tau}_j - x_j)^2 \,|\, \tilde{\mu}, \tilde{s}_j\big]$. By Lemma 2, this individual chooses

$$x_j = \mathrm{E}\big[\tilde{\tau}_j \mid \tilde{\mu}, \tilde{s}_j\big] = \tilde{\mu} + \bar{e}_\lambda \mathrm{E}\big[\tilde{\tau}_j - \tilde{\mu} \mid \tilde{\mu}, \tilde{s}_j = \tilde{\tau}_i\big] + (1-\bar{e}_\lambda)\underbrace{\mathrm{E}\big[\tilde{\tau}_j - \tilde{\mu} \mid \tilde{\mu}\big]}_{=0} = \tilde{\mu} + \bar{e}_\lambda\rho(\tilde{s}_j - \tilde{\mu}),$$

where the last equality is obtained using the projection theorem. ■

Proof of Lemma 5

Suppose first that there is no default option. From the proof of Proposition 3, we know that the welfare of any one skilled individual $i \in I_\lambda$ is given by

$$W_{\mathrm{N},i} = -(1-e_i)(\Sigma_\mu + \Sigma) - \frac{c}{2}e_i^2 + \frac{(1-\lambda)p}{\lambda}.$$

The welfare of any one unskilled individual $i \in I \setminus I_\lambda$ is given by

$$W_{\mathrm{N},i} = -(\Sigma_\mu + \Sigma) + v_{\mathrm{N}} - p,$$

and so total welfare is

$$\begin{aligned} W_{\mathrm{N}} \equiv \int_I W_{\mathrm{N},i}\, d\gamma &= \int_{I_\lambda} \Big[-(1-e_i)(\Sigma_\mu + \Sigma) - \frac{c}{2}e_i^2\Big] d\gamma + \int_{I \setminus I_\lambda} \Big[-(\Sigma_\mu + \Sigma) + v_{\mathrm{N}}\Big] d\gamma \\ &= -(\Sigma_\mu + \Sigma) + \int_{I_\lambda} \Big[e_i(\Sigma_\mu + \Sigma) - \frac{c}{2}e_i^2\Big] d\gamma + (1-\lambda)v_{\mathrm{N}}. \end{aligned}$$

In equilibrium, we know from Proposition 3 that $e_i = \bar{e}_\lambda = \frac{\Sigma_\mu + \Sigma}{c}$, $p = \theta v_{\mathrm{N}}$, and $v_{\mathrm{N}} = \frac{(\Sigma_\mu + \rho\Sigma)^2}{\Sigma_\mu + \Sigma}\bar{e}_\lambda^2$. After using these expressions in the total welfare function above, we get

$$W_{\mathrm{N}} = -(\Sigma_\mu + \Sigma) + \lambda\Big[\bar{e}_\lambda(\Sigma_\mu + \Sigma) - \frac{c}{2}\bar{e}_\lambda^2\Big] + (1-\lambda)\frac{(\Sigma_\mu + \rho\Sigma)^2}{\Sigma_\mu + \Sigma}\bar{e}_\lambda^2,$$

which simplifies to (14). The calculations are similar with the default option. ■

Proof of Proposition 5

Manipulations of (14) and (15) show that $W_{\mathrm{N}} > W_{\mathrm{D}}$ if and only if

$$-c^2\Sigma_\mu + c\frac{\lambda}{2}\Sigma_\mu(\Sigma_\mu + 2\Sigma) + (1-\lambda)\Big[(\Sigma_\mu + \rho\Sigma)^2(\Sigma_\mu + 2\Sigma) - \rho^2\Sigma^3\Big] > 0. \tag{A10}$$

Since the left-hand-side of this inequality is quadratic in c, positive at $c = 0$, and negative for large c, the inequality holds if and only if

$$c < \frac{\lambda}{4}(\Sigma_\mu + 2\Sigma) + \frac{1}{2\Sigma_\mu}\sqrt{\frac{\lambda^2}{4}(\Sigma_\mu + 2\Sigma)^2 + 4(1-\lambda)\Sigma_\mu\Big[(\Sigma_\mu + \rho\Sigma)^2(\Sigma_\mu + 2\Sigma) - \rho^2\Sigma^3\Big]}.$$

Since c must be larger than $\Sigma_\mu + \Sigma$ by assumption, it must be the case that this upper bound for c is larger than $\Sigma_\mu + \Sigma$ for $W_{\mathrm{N}} > W_{\mathrm{D}}$ to ever be possible. Equivalently, this will be the case when (A10) evaluated at $c = \Sigma_\mu + \Sigma$ is greater than zero. Straightforward calculations show that this inequality simplifies to (18). $\blacksquare$

References

Ahdieh, Robert. 2011. "Beyond Individualism in Law and Economics." *Boston University Law Review*, 91: 43-85.

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2008. "When Does Coordination Require Centralization?" *American Economic Review*, 98: 145-179.

Angeletos, George-Marios, and Alessandro Pavan. 2007. "Efficient Use of Information and Social Value of Information." *Econometrica*, 75: 1103-1142.

Arrow, Kenneth. 1994. "Methodological Individualism and Social Knowledge." *American Economic Review*, 84: 1-9.

Baker, Tom, and Timothy Lytton. 2010. "Allowing Patients to Waive the Right to Sue for Medical Malpractice: A Response to Thaler and Sunstein." *Northwestern University Law Review*, 104: 233-252.

Beshears, John, James Choi, David Laibson, Brigitte Madrian, and Katherine Milkman. 2010. "The Effect of Providing Peer Information on Retirement Savings Decisions." National Bureau of Economic Research, Working Paper.

Bolton, Patrick, and Joseph Farrell. 1990. "Decentralization, Duplication, and Delay." *Journal of Political Economy*, 98: 803-826.

Brown, Christina, and Aradhna Krishna. 2004. "The Skeptical Shopper: A Metacognitive Account for the Effects of Default Options on Choice." *Journal of Consumer Research*, 31: 529-539.

Brown, Jeffrey, Zoran Ivković, Paul Smith, and Scott Weisbenner. 2008. "Neighbors Matter: Causal Community Effects and Stock Market Participation." *Journal of Finance*, 63: 1509-1531.

Burguet, Roberto, and Xavier Vives. 2000. "Social Learning and Costly Information Acquisition." *Economic Theory*, 15: 185-205.

Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin. 2003. "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism'." *University of Pennsylvania Law Review*, 151: 1211-1254.

Carlin, Bruce. 2009. Strategic Price Complexity in Retail Financial Markets. *Journal of Financial Economics*, 91: 278-287.

Carlin, Bruce, and Gustavo Manso. 2011. Obfuscation, Learning, and the Evolution of Investor Sophistication. *Review of Financial Studies*, 24: 754-785.

Carroll, Gabriel, James Choi, David Laibson, Brigitte Madrian, and Andrew Metrick. 2009. "Optimal Defaults and Active Decisions." *Quarterly Journal of Economics*, 124: 1639-1674.

Choi, James, David Laibson, Brigitte Madrian, and Andrew Metrick. 2002. "Defined Contribution Pensions: Plan Rules, Participant Choices, and the Path of Least Resistance." *Tax Policy and the Economy Volume*, 16: 67-113.

Choi, James, David Laibson, Brigitte Madrian, and Andrew Metrick. 2003. "Optimal Defaults." *American Economic Review*, 93: 180-185.

Duffie, Darrell, and Gustavo Manso. 2007. "Information Percolation in Large Markets." *American Economic Review*, 97: 203-209.

Duflo, Esther, and Emmanuel Saez. 2002. "Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues' Choices." *Journal of Public Economics*, 85: 121-148.

Duflo, Esther, and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics*, 118: 815-842.

Ellison, Glenn, and Drew Fudenberg. 1993. "Rules of Thumb for Social Learning." *Journal of Political Economy*, 101: 612-643.

Ellison, Glenn, and Drew Fudenberg. 1995. "Word-of-Mouth Communication and Social Learning." *Quarterly Journal of Economics*, 110: 93-125.

Glaeser, Edward. 1999. "Learning in Cities." *Journal of Urban Economics*, 46: 254-277.

Glaeser, Edward. 2006. "Paternalism and Psychology." *University of Chicago Law Review*, 73: 133-156.

Glaeser, Edward, and David Maré. 2001. "Cities and Skills." *Journal of Labor Economics*, 19: 316-342.

Glaeser, Edward, Bruce Sacerdote, and José Scheinkman. 2003. "The Social Multiplier." *Journal of the European Economic Association*, 1: 345-353.

Grossman, Sanford. 1976. "On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information." *Journal of Finance*, 31: 573-585.

Grossman, Sanford, and Joseph Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review*, 70: 393-408.

Hayek, Friedrich. 1945. "The Use of Knowledge in Society." *American Economic Review*, 35: 519-530.

Hirshleifer, Jack. 1971. "The Private and Social Value of Information and the Reward to Inventive Activity." *American Economic Review*, 61: 561-574.

Hirshleifer, Jack. 1973. "Where Are We in the Theory of Information?" *American Economic Review*, 63: 31-39.

Hong, Harrison, Jeffrey Kubik, and Jeremy Stein. 2004. "Social Interaction and Stock-Market Participation." *Journal of Finance*, 59: 137-163.

Johnson, Eric, and Daniel Goldstein. 2003. "Do Defaults Save Lives?" *Science*, 302: 1338-1339.

Kaustia, Markku, and Samuli Knüpfer. 2012. "Peer Performance and Stock Market Entry." *Journal of Financial Economics*, 104: 321-338.

Korobkin, Russell. 2009. "Libertarian Welfarism." *California Law Review*, 97: 1651-1685.

Lange, Oskar. 1938. *On the Economic Theory of Socialism.* University of Minnesota Press, Minneapolis.

Lerner, Abba. 1938. "Theory and Practice in Socialist Economics." *Review of Economic Studies*, 6: 71-75.

Lusardi, Annamaria, and Olivia Mitchell. 2007. "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth." *Journal of Monetary Economics*, 54: 205-224.

Madrian, Brigitte, and Dennis Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics*, 116: 1149-1187.

Manski, Charles. 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives*, 14: 115-136.

Manski, Charles. 2004. "Social Learning from Private Experiences: The Dynamics of the Selection Problem." *Review of Economic Studies*, 71: 443-458.

McKenzie, Craig, Michael Liersch, and Stacey Finkelstein. 2006. "Recommendations Implicit in Policy Defaults." *Psychological Science*, 17: 414-420.

Mitchell, Gregory. 2005. "Libertarian Paternalism Is an Oxymoron," *Northwestern Law Review*, 99: 1245-1277.

Morris, Stephen, and Hyun Song Shin. 2002. "Social Value of Public Information." *American Economic Review*, 92: 1521-1534.

Rostbøll, Christian. 2005. "Preferences and Paternalism on Freedom and Deliberative Democracy." *Political Theory*, 33: 370-396.

Shiller, Robert, and John Pound. 1989. "Survey Evidence on Diffusion of Interest and Information Among Investors." *Journal of Economic Behavior and Organization*, 121: 47-66.

Sorensen, Alan. 2006. "Social Learning and Health Plan Choice." *RAND Journal of Economics*, 37: 929-945.

Stiglitz, Joseph. 1994. *Whither Socialism?* MIT Press, Cambridge.

Stiglitz, Joseph 2000. "The Contributions of the Economics of Information to Twentieth Century Economics." *Quarterly Journal of Economics*, 115: 1441-1478.

Taylor, Fred. 1929. "The Guidance of Production in a Socialist State." *American Economic Review*, 19: 1-8.

Thaler, Richard, and Shlomo Benartzi. 2004. "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy*, 112: S164-S187.

Thaler, Richard, and Cass Sunstein. 2003. "Libertarian Paternalism." *American Economic Review*, 93: 175-179.

Thaler, Richard, and Cass Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press, New Haven.

Whitman, Douglas, and Mario Rizzo. 2007. "Paternalist Slopes." *NYU Journal of Law and Liberty*, 2: 411-443.