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

Bruce Ian Carlin†
Simon Gervais‡
Gustavo Manso§

March 12, 2013

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 have greatly benefitted from comments and suggestions from the Editor (David Hirshleifer) and two anonymous referees. We would also 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.

Libertarian Paternalism, Information Production, and Financial Decision-Making

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.

1 Introduction

Financial sophistication has lagged behind the growing complexity of retail markets (e.g., NASD Literacy Survey, 2003).[1] What to do about this disparity remains hotly debated. Whereas some investigators 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. 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 is a compromise between government intervention and free markets, whereby centralized and decentralized uses of information can coexist to maximize welfare. A 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. However, just as 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 and the net effect of default options on total welfare. Our analysis is grounded in the idea that default options provide information to market participants, and that this information reduces individuals' willingness to educate themselves about the choices available to them. Ultimately, this changes the financial decisions that individuals make. A good example of this can be found in the empirical work of Madrian and Shea (2001). They find that employees hired prior to the addition of default options in their company's 401(k) plan tend to adopt these defaults when they join the plan even though

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

they are not subject to them. 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).

Decreasing the incentives to produce information is particularly costly because individuals frequently share the information they gather with their peers. For example, Duflo and Saez (2003) show that individuals' decision to participate in a tax-advantageous retirement plan is highly correlated with that of colleagues who have independently been advised to do so. Similarly, Duflo and Saez (2002), Sorensen (2006), and Beshears et al. (2012) document that individuals learn about their economic decisions through their interactions with each other.[2] It is this combination of social learning and reduced information-gathering incentives induced by defaults that is the focus of our paper.

Our analysis considers both a setting in which information spreads 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).[3] In both settings, 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 individuals in the economy. She must decide between two policies: (i) institute a default option that implicitly reveals useful information to individuals; (ii) let individuals make their own choices without guidance from an informative default. Individuals can exert costly effort to find out more about their own type, and people may also become informed through social interactions or by contacting skilled agents.

We derive conditions under which default options are optimal and describe when they destroy social surplus. The tradeoff pertains to the fact that the information contained in the default option provided by the social planner reduces individuals' incentive to gather and share any additional

[2]Indeed, social interactions have been shown to affect a variety of financial 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., 2012), and to buy stocks (Shiller and Pound, 1989). For a more general survey of the literature on social interactions, see Manski (2000).

[3]In the advice market, individuals learn to make better decisions by interacting with their skilled peers. As such, 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.

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, the individual-specific information is more valuable, and/or the planner's information is imprecise, providing a default option reduces welfare. Under these conditions, a social planner maximizes welfare by letting market participants fend for themselves and allowing social learning to thrive.

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 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.

2 Context and Related Literature

As mentioned in the introduction, our model combines two economic forces that have been empirically documented in various contexts: (i) individuals learn more efficiently about the important economic decisions that they face when they can interact with each other; (ii) default options reduce the need for individuals to analyze the choices that are available to them. This choice of forces serves to highlight the informational tradeoff that is inherently part of default options. In this section, we discuss how our study of this tradeoff adds to the recent debates about libertarian paternalism, to existing models of default options, and to a long-standing strand of economic literature about the social value of information.

2.1 The Pros and Cons of Libertarian Paternalism

In a seminal article, Jolls et al. (1998) propose that law and economics adopt an approach that is better grounded in human behavior. In essence, their approach prompts lawmakers to internalize the biases that are known to systematically affect individuals' decisions. Adopting this behavioral approach, Camerer et al. (2003), and Thaler and Sunstein (2003) suggest soft versions of paternalism for policymaking. In the former, the authors advocate the use of paternalism in contexts where it is greatly beneficial to those who make mistakes but has little effect on others. The latter introduces libertarian paternalism, in which agents are provided with options that can guide them but that they are free to ignore, an approach expanded upon by Thaler and Sunstein (2008).[4]

Since then, soft paternalism has penetrated the realms of policymaking in several different contexts, from credit cards (Barr et al., 2008b) to mortgage lending (Barr et al., 2008a, 2008b), from private retirement plans (Choi et al., 2004) to social security (Cronqvist and Thaler, 2004), and from organ donation (Johnson and Goldstein, 2003) to healthy living (Loewenstein et al., 2007). Some proponents of such policies argue that they protect consumers not only from their own mistakes but also from the exploitation of their behavioral biases by businesses (e.g., Bar-Gill and Warren, 2008; Barr et al., 2008b).

These policies are not without their detractors. For example, Glaeser (2006) suggests that in some contexts libertarian paternalism may be hard to publicly monitor and may lead to hard paternalism.[5] He also warns about the possibility that social planners are not immune from making errors or having biases, which may affect the value of default options. Mitchell (2005) questions the redistributive consequences of libertarian paternalism. Korobkin (2009) argues that, because

[4]Thaler and Sunstein (2008), and Thaler et al. (2010) add choice architecture as a tool of libertarian paternalism to non-forcefully guide people away from the bad choices induced by their behavioral tendencies.

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

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. 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. More generally, Wright and Ginsburg (2012) argue that policies based on libertarian paternalism may have been prematurely implemented as the welfare tradeoffs have yet to be properly identified and measured.

Our theory highlights one such cost, namely the negative effect that defaults have on the production of information that improves individual decisions. Specifically, our paper adds agent heterogeneity and (social) learning to the list of factors that potentially reduce the benefits of libertarian paternalism. Rachlinski (2006) argues against any form of paternalism when the heterogeneity across agents is large. Likewise, Rizzo and Whitman (2009) argue that, to be effective, paternalistic policymakers must account for the heterogeneity in the population that they seek to protect. Our work shows that indeed libertarian paternalism is least useful, and even potentially harmful, when individuals differ greatly from each other. The same authors argue that new paternalists must overcome significant knowledge-based obstacles to make libertarian paternalism successful. For example, they write (page 967): "Even knowing that the average or typical person is in need of paternalistic assistance is not sufficient because... the average or typical person might respond in counterproductive ways, such as reducing self-corrective efforts." The effort that individuals must exert in our model in order to learn about the decisions they face is consistent with this view. Our model further shows that reductions in self-corrective efforts are especially damageable when information can be shared.

Our work ultimately leads to the conclusion that libertarian paternalism should be used judiciously rather than as a blanket policy. Indeed, 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.[6]

[6]Similarly, Ahdieh (2011) stresses the importance for any public intervention aimed at individuals to internalize the social dynamics that it may affect.

2.2 Models of Defaults

To our knowledge, the effects of information sharing on the welfare of default options have not been formally analyzed before. However, some models have been proposed to deal with other aspects of defaults options. Choi et al. (2003) develop a model that analyzes 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. In an extension to this model, Carroll et al. (2009) add the possibility for the social planner to force agents into a decision, as opposed to letting them navigate through defaults. Although such "active decisions" (as they term them) increase the overall effort costs of agents, their merit is to prompt agents into action which, as the authors show, is valuable when the degree of procrastination is extreme.

Our model differs from these other models in several respects. First, although our model could accommodate the procrastination bias that they assume, we restrict our analysis to fully rational agents.[7] Second, whereas these other models involve only one agent and one planner, ours highlights the fact that defaults are set for a group of agents whose preferences and optimal decisions are in general heterogeneous. Third, our model emphasizes the informational aspect of default options, which is not considered in these papers. Finally, the social planner in these models motivates agents to make active decisions by selecting defaults that are biased and that end up hurting those who stick with them. In contrast, our social planner always picks an unbiased default when she announces one, protecting those who ultimately end up in that default.

The informational aspect of default options is also considered in two recent papers. An experiment by Phatak (2012) shows that, when the data to set defaults must come from actual decisions, the fact that these decisions are themselves affected by the defaults makes subsequent planner intervention less valuable. Indeed, by reducing information-gathering incentives, current defaults render future defaults less useful since they incorporate no new information. Closer to our paper is

[7] A previous version of this paper, available from the authors upon request, included a procrastination component in which agents mistakenly overestimate the effort cost required at the outset in order to improve their decisions and future outcomes. Because this additional feature of the model did not affect the informational aspect of the results, the main purpose of our analysis, we elected to remove it.

the work by Caplin and Martin (2012) who show, both theoretically and experimentally, how informative defaults reduce the investment in attention made by agents who face an economic decision, a phenomenon they call the "drop out effect."

As in these papers, our model shows that the presence of a default reduces the investment in effort that agents are willing to make in order to improve the quality of their decisions. In contrast to the former, our model focuses on the production and exchange of information by market participants and shows how defaults affect them. In contrast to the latter, we do not assume that agents misinterpret the usefulness of the defaults that are chosen by the planner, and so our welfare results derive from better coordination, not from debiasing.

Existing theories of defaults also differ in their treatment of the status quo, or what happens in the absence of a default. In fact, Choi et al. (2003) completely avoid the issue by requiring the social planner to always create a default. Carroll et al. (2009) add the possibility for the social planner to effectively force individuals into action by setting a default that is extremely costly to them. Whereas Caplin and Martin (2012) assume that inattentive agents end up with a randomly drawn decision when the social planner does not set a default, we assume that the status quo decision of agents is pre-set at a given (constant) value. In both cases, however, the absence of a default is correctly interpreted by agents as a commitment from the government not to assist agents in their choices; that is, the absence of a default is assumed to be fully non-paternalistic. Although this commitment is not formalized in either paper, one could imagine that agents update about the social planner's intentions based upon the public expense that she incurs to gather the information needed to produce a useful default.

These modeling decisions are not completely innocuous as, in some settings, the status quo could be deliberately allowed to prevail by the social planner and thus implicitly represents some form of default in and of itself. For example, consider a savings decision in which people have to decide what fraction of their income to invest for retirement. Without a default, the status quo is zero savings, which people may view as the recommended option. That is, individuals may come to view decisions that are not disallowed as good ones.[8] In contrast, for many financial decisions, there is no

[8]Note that, if agents do not know the intention of the social planner for every economic decision they face, the

information to be learned from the absence of a default option proactively set by the social planner. For example, consider the proposal to make the 30-year fixed rate mortgage the default option in the real estate market (Barr et al., 2008a). Offering such a default conveys a recommendation to home buyers. When the default is not imposed, home buyers must choose among a menu of options, none of which is preferred by a social planner. In such case, no information is communicated to market participants. The same case may be made regarding proposals for a default "clean" credit card (Warren, 2007; Barr et al., 2008b), a default portfolio allocation, or a default life annuity.

In our model, because agents are rational and because they can make economic decisions for free, these considerations do not have any bite. Indeed, the uninformative defaults that we assume are isomorphic to random defaults, the absence of a default, and active decisions. In all cases, agents internalize all of the information that is available to them and make the optimal choice based on this information.

2.3 The Social Value of Public Information

Although our model is designed specifically to match the context of default options, the effects of public information on welfare that it yields are not specific to it. Indeed, starting with Hirshleifer (1971), a long line of papers questions the welfare benefits of making information publicly available in various contexts. For example, Bikhchandani et al. (1992) show that public disclosures can engender inefficient informational cascades when they lead agents to ignore the value of their private information. Teoh (1997) studies the provision of public goods and shows that public disclosures may exacerbate the underinvestment and free-riding problems associated with team production. Burguet and Vives (2000) model a series of short-lived agents who may exert costly effort to learn about a common random variable. Since agents fail to account for the information their actions reveal to later generations, the release of public information may decrease welfare since agents invest less in acquiring private information. Morris and Shin (2002) study a beauty contest game and show that the release of public information may induce agents to ignore their private

status quo leads them to revise downward their beliefs that she actually learned new, useful information. Thus, real-life situations may be somewhere in between this example and our model.

signals and inefficiently herd. Angeletos and Pavan (2007) generalize Morris and Shin's (2002) analysis and characterize the conditions under which public disclosure destroys economic surplus, given that agent actions may be substitutes or complements. Finally, Amador and Weill (2012) consider a continuous-time model to explore how public disclosures inefficiently slow learning.

In addition to linking libertarian paternalism to this line of research, our paper contributes to the literature on public disclosure in several respects. First, in contrast to previous studies, agents in our setting are heterogeneous. As we show, this heterogeneity in people's needs or attributes can have an important impact on the value of public disclosure. Specifically, a default that induces all agents to converge to a similar decision can be detrimental when their needs are disperse. For example, it may be suboptimal for a default to lead the majority of participants in a 401(k) plan to adopt a savings rate of 10% if the optimal savings rate for many of them is close to zero or 20%. Indeed, this is the very problem that Tergesen (2011) exposes in a Wall Street Journal study on the use of default options in retirement plans.[9] Further, since agent heterogeneity increases with the magnitude of the financial decision they take, it is also the case that the social planner should refrain from issuing a default when stakes are larger. This provides another novel intuition that is not present in the extant literature.

Second, although the aforementioned papers all rely on some economic externality for their results, the externality that is key to our model of defaults comes from the joint information acquisition process of agents. Specifically, the externality originates from the fact that an agent's learning is a complement to that of other agents. In particular, our model does not rely on beauty contests (Morris and Shin, 2002), generational concerns (Burguet and Vives, 2000), or a public goods problem (Teoh, 1997). Finally, our model differs from existing models in that the externality is endogenized through an advice channel that allows for information sales. How public disclosures affect this advice channel is new to the literature.

[9] The article quotes Brigitte Madrian as saying: "Automatic enrollment is a double-edged sword. On the one hand, there's more participation. On the other hand, lots of employees are stuck at whatever default the employer selects."

3 Libertarian Paternalism with Social Learning

3.1 The Model's Setup

Consider an economy that 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 of an 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. 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 $i \neq j$.

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

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

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.[10]

[10]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.

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 agents to learn about the decision that they have to make. As such, the parameter α measures the degree of this information externality.

While we use this reduced-form model for parsimony, it accommodates simple micro-foundations. For example, assume individuals exert effort to find a source of information that allows all individuals to learn their type, like an insightful article, a useful website, or a trustworthy financial advisor. Moreover, after choosing an effort level e_i, each individual i has a probability α of meeting some other individual randomly drawn from the population. When two individuals meet, they can avoid the duplication of their effort, and thus can find a source of information that is helpful to both with probability $e_i + e_j$. On the other hand, an individual i who does not meet anyone else finds such an information source with probability e_i only. Ex ante, then, each individual observes his type with the probability in (3).[11]

This information structure is well-suited for many empirical settings in which libertarian paternalism is applied. For example, consider a firm's employees who face a 401(k) allocation problem. Even though they may have different needs because of underlying demographic factors, the effort that one employee exerts can spill over to the success that others have in determining their optimal asset allocation, as employees share their findings with each other. This last consideration is in fact explored by Duflo and Saez (2003), who analyze a randomized experiment in which a subset

[11] With this micro-foundation, individuals choose their effort level before knowing whether they will meet another individual. The analysis would be unchanged if individuals made this decision after a potential meeting, since it is easy to show that optimal effort level would not depend on whether or not a meeting takes place.

of a university's employees were encouraged to attend an educational event about enrolling in a Tax Deferred Account (TDA) retirement plan. Individuals who received that encouragement participated at a significantly higher rate in the TDA's, compared to the control group. Surprisingly, however, in departments that were treated, participation was almost as high for employees who were not specifically encouraged, a clear product of social interactions and information spillovers.

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 individuals make.

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 effectively reveals $\tilde{s}$ 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 individuals, as empirically documented by Madrian and Shea (2001). Since agents 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.

3.2 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 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 people 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 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.[12] When there is a default option and the individual does not 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 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 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, individual* i *chooses his effort level* e_i *to maximize*

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

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

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

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

[12] 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.

variance reduction. Importantly, in both scenarios, individual i fails to internalize the externality that his effort creates. That is, because e_i is infinitesimally small in $\bar{e}$, the size of α does not affect individual i's choice of effort, leaving equilibrium effort levels below their first-best value for all $\alpha > 0$. As such, the social planner's decision to offer a default depends on how important these deviations are in the two scenarios.

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. The following proposition characterizes each individual's effort choice, with and without a default option.

Proposition 1. *If the social planner does not adopt a default option, each 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 $\bar{e} = e^{\mathrm{N}}$. An individual i who observes a fully informative signal opts out of the default option and chooses $x_i = \tilde{\tau}_i$. An individual i who does not become informed chooses $x_i = 0$.

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

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

where $\delta = \Sigma_\mu/(\Sigma_\mu + \Sigma_\epsilon)$, and the average effort level of the population is $\bar{e} = e^{\mathrm{D}}$. An 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 all individuals exert the same effort, $\bar{e}$ is equal to e^{N} without a default and to e^{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 each 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). This effect is also consistent with the work of Brown et al. (2011) who document that participants in a retirement plan who adopt the plan's defaults do so in part because they lack the information necessary to do otherwise.

Table 1 abstracts from one additional force that makes the availability of a default option advantageous, namely the fact that the overall cost of information production is greater without a default option, as individuals exert greater effort to produce it. The following proposition takes this additional tradeoff into account to derive and compare the total welfare with and without a default option.

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

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

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:*

$$2(\Sigma_\mu + \Sigma) < c < \left(\frac{1}{2} + \alpha\right)\left[(2-\delta)\Sigma_\mu + 2\Sigma\right]. \tag{8}$$

This region is non-empty if and only if

$$\frac{\delta\Sigma_\mu}{\Sigma_\mu + \Sigma} < 2 - \frac{2}{\left(\frac{1}{2} + \alpha\right)}. \tag{9}$$

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).[13] This arises because the presence of a default option reduces people's incentives to learn about the problem they face, which in turn makes information-sharing less effective in the economy. Specifically, since the right-hand sides of (8) and (9) are increasing in α, it is better for the social planner to leave the production of knowledge to individuals when information is easy to communicate to others (α greater than $\frac{1}{2}$ and large). 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 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, 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

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

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.

4 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 3. 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 3 to a context in which some individuals can (and will) seek the advice of other individuals in the economy. More specifically, we assume that the population consists of skilled individuals (fraction λ) and unskilled individuals (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 $C(e_i) = \frac{ce_i^2}{2}$, individual $i \in I_\lambda$ receives a signal that reveals his type $\tilde{\tau}_i$ with probability e_i.

The other individuals $j \in I \backslash I_\lambda$ are unskilled in that gathering information about their own type is prohibitively costly. However, these unskilled individuals are allowed to purchase information from one randomly-picked skilled individual and to rationally use this information to make their financial decision x_j.[14] 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.[15]

Because all information sales happen at the same time and because skilled individuals are identified as such, we implicitly assume that unskilled individuals cannot resell their purchased signal. This makes sense as our main intention is to capture the idea that some individuals are seen to have the ability and technology to learn about the problem at hand and this innately turns them into advisors.[16]

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*

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

[14]Note that a skilled individual may end up getting picked by several unskilled individuals. That is, the random selection is done with replacement.

[15]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 a 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.

[16]Note also that this implicit assumption would arise endogenously even if we assumed that the information gathered by skilled individuals percolates across the population through multiple rounds of trading. Indeed, if individuals can only meet one peer, they always prefer consulting one known to be skilled, as in our model, since unskilled individuals may have yet to meet an informed counter-party and so are more likely to be selling noise.

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 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, Γ 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 $v_{\mathrm{N}} = \left[\Gamma + \rho(1-\Gamma)\right]^2 \Sigma_\tau \bar{e}_\lambda^2$ and $v_{\mathrm{D}} = \rho^2(1-\Gamma)\Sigma_\tau \bar{e}_\lambda^2$. 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 $v_{\mathrm{N}} > v_{\mathrm{D}}$ 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 $\sigma = \mathrm{D}$ if a default option is made available (and $\sigma = \mathrm{N}$ otherwise). When $\theta = 1$ ($\theta = 0$), the skilled (unskilled) individual extracts all the surplus from the transaction.[17] 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.[18] We start with the following result, which describes the equilibrium in the absence of a default option.

Proposition 3. *Suppose that the social planner does not adopt a default option.*
(i) 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 his own information allowed him to learn his type $\tilde{\tau}_i$.
(ii) Also, 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 makes an economic decision x_j that weighs this signal using Bayes' rule $\big($factor $\frac{\Sigma_\mu + \rho\Sigma}{\Sigma_\mu + \Sigma}\big)$ and the probability $\bar{e}_\lambda$ that his skilled counter-party was informed:

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

The skilled individuals' behavior is the same as in Section 3. 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

[17] 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.

[18] 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.

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. *Suppose that the social planner adopts a default option.*
(i) 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 his own information allowed him to learn his type $\tilde{\tau}_i$.
(ii) Also, 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 $p = \theta v_{\mathrm{D}}$, with v_{D} given by (11), and adjusts his economic decision x_j away from the default ($\tilde{\mu}$) according to the correlation across types (ρ) and the probability $\bar{e}_\lambda$ that his skilled counter-party was informed:

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

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 3, 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 3, 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.[19] That is, unskilled individuals 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 3. 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.[20]

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

[19]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.

[20]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).

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.

5 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. 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 and in turn may alter the success of information aggregation which, as suggested by Duflo and Saez (2003), is often facilitated by social learning or 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.

Thaler and Sunstein (2008) write that a nudge is "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives" (page 6). As we show, however, informative defaults do change people's economic incentives to gather information, and this can be socially costly. In this way, 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

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

This completes the proof. ■

Proof of Proposition 1

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

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

The first-order condition for this problem is

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

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

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

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

The first-order condition for this problem is

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

which leads to (7) and to $\bar{e} \equiv \int_I e_i d\gamma = e^{\mathrm{D}}$. 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 individuals without a default option,

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

and with a default option,

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

A simple comparison of (A1) and (A2) yields the second inequality in (8). The first inequality in (8) is by assumption. The region is non-empty if and only if

$$2(\Sigma_\mu + \Sigma) < \left(\frac{1}{2} + \alpha\right)\big[(2-\delta)\Sigma_\mu + 2\Sigma\big],$$

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

$$\begin{aligned}\mathrm{E}\Big[\tilde{U}_i(x_j)-p\Big] &= \mathrm{E}\Big[(\tilde{\tau}_j-\bar{e}_\lambda\beta_j\tilde{s}_j)^2\Big]-p \\ &= \bar{e}_\lambda\mathrm{E}\Big[(\tilde{\tau}_j-\bar{e}_\lambda\beta_j\tilde{\tau}_i)^2\Big]+(1-\bar{e}_\lambda)\mathrm{E}\Big[(\tilde{\tau}_j-\bar{e}_\lambda\beta_j\tilde{\eta}_i)^2\Big]-p, \qquad \text{(A3)}\end{aligned}$$

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 (A3) 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

$$\mathrm{E}\Big[\tilde{U}_i(x_j)-p\Big]=\Sigma_\mu+\Sigma-\bar{e}_\lambda^2\frac{(\Sigma_\mu+\rho\Sigma)^2}{\Sigma_\mu+\Sigma}-p. \qquad \text{(A4)}$$

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

$$\mathrm{E}\Big[\tilde{U}_i(x_j)\Big]=\mathrm{E}\Big[\tilde{\tau}_j^2\Big]=\Sigma_\mu+\Sigma. \qquad \text{(A5)}$$

Thus the largest price p that makes individual j indifferent between purchasing and not purchasing $\tilde{s}_j$ is that which makes (A4) and (A5) 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

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

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

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

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

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

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

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

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

$$-c^2\Sigma_\mu + c\frac{\lambda}{2}\Sigma_\mu(\Sigma_\mu + 2\Sigma) + (1-\lambda)\Big[(\Sigma_\mu + \rho\Sigma)^2(\Sigma_\mu + 2\Sigma) - \rho^2\Sigma^3\Big] > 0. \tag{A6}$$

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 (A6) evaluated at $c = \Sigma_\mu + \Sigma$ is greater than zero. Straightforward calculations show that this inequality simplifies to (18). ■

References

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

Amador, Manuel, and Pierre-Olivier Weill. 2012. "Learning from Private and Public Observations of Others' Actions." *Journal of Economic Theory*, 147: 910-940.

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

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.

Bar-Gill, Oren, and Elizabeth Warren. 2008. "Making Credit Safer." *University of Pennsylvania Law Review*, 157: 1-101.

Barr, Michael, Sendhil Mullainathan, and Eldar Shafir. 2008a. "An Opt-Out Home Mortgage System." The Brookings Institution, Hamilton Project Discussion Paper 2008-14.

Barr, Michael, Sendhil Mullainathan, and Eldar Shafir. 2008b. "Behaviorally Informed Financial Services Regulation." New America Foundation, White Paper.

Beshears, John, James Choi, David Laibson, Brigitte Madrian, and Katherine Milkman. 2012. "The Effect of Providing Peer Information on Retirement Savings Decisions." Stanford University, Working Paper.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." *Journal of Political Economy*, 100: 992-1026.

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, Anne Farrell, and Scott Weisbenner. 2011. "The Downside of Defaults." University of Illinois at Urbana-Champaign, Working Paper.

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.

Caplin, Andrew, and Daniel Martin. 2012. "Defaults and Attention: The Drop Out Effect." New York University, Working Paper.

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.

Choi, James, David Laibson, Brigitte Madrian, and Andrew Metrick. 2004. "For Better or for Worse Default Effects and 401(K) Savings Behavior." In *Perspectives on the Economics of Aging*, D. A. Wise (Ed.), 81-125. University of Chicago Press, Chicago.

Cronqvist, Henrik, and Richard Thaler. 2004. "Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience." *American Economic Review*, 94: 424-428.

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.

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.

Jolls, Christine, Cass Sunstein, and Richard Thaler. 1998. "A Behavioral Approach to Law and Economics." *Stanford Law Review*, 50: 1471-1550.

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.

Loewenstein, George, Troyen Brennan, and Kevin Volpp. 2007. "Asymmetric Paternalism to Improve Health Behaviors." *Journal of the American Medical Association*, 298: 2415-2417.

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.

Phatak, Narahari. 2012. "Menu-Based Complexity: Experiments on Choice Over Lotteries." University of California at Berkeley, Working Paper.

Rachlinski, Jeffrey. 2006. "Cognitive Errors, Individual Differences, and Paternalism." *University of Chicago Law Review*, 73: 207-229.

Rizzo, Mario, and Douglas Whitman. 2009. "The Knowledge Problem of New Paternalism." *BYU Law Review*, 2009: 906-968.

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.

Tergeson, Anne. 2011. "401(k) Law Suppresses Saving for Retirement." *Wall Street Journal*, July 7.

Teoh, Siew Hong. 1997. "Information Disclosure and Voluntary Contributions to Public Goods." *RAND Journal of Economics*, 28: 385-406.

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.

Thaler, Richard, Cass Sunstein, and John Balz. 2010. "Choice Architecture." University of Chicago, Working Paper.

Warren, Elizabeth. 2007. "Unsafe at any Rate." *Democracy: A Journal of Ideas*, 5: 8-19.

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

Wright, Joshua, and Douglas Ginsburg. 2012. "Behavioral Law and Economics: Its Origins, Fatal Flaws, and Implications for Liberty." *Northwestern University Law Review*, 106: 1033-1088.