NBER WORKING PAPER SERIES

WHEN DOES LIBERTARIAN PATERNALISM WORK?

Bruce Ian Carlin
Simon Gervais
Gustavo Manso

Working Paper 15139
http://www.nber.org/papers/w15139

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2009

When Does Libertarian Paternalism Work?
Bruce Ian Carlin, Simon Gervais, and Gustavo Manso
NBER Working Paper No. 15139
July 2009
JEL No. G18,G38,H11

ABSTRACT

We develop a theoretical model to study the effects of libertarian paternalism on knowledge acquisition and social learning. Individuals in our model are permitted to appreciate and use the information content in the default options set by the government. We show that in some settings libertarian paternalism may decrease welfare because default options slow information aggregation in the market. We also analyze what happens when the government acquires imprecise information about individuals, and characterize its incentives to avoid full disclosure of its information to the market, even when it has perfect information. Finally, we consider a market in which individuals can sell their information to others and show that the presence of default options causes the quality of advice to decrease, which may lower social welfare.

Bruce Ian Carlin
Anderson Graduate School of Management
UCLA
110 Westwood Plaza, Suite C519
Los Angeles, CA 90095-1481
and NBER
bruce.carlin@anderson.ucla.edu

Simon Gervais
Fuqua School of Business
Duke University
One Towerview Drive
Durham, NC 27708-0120
sgervais@duke.edu

Gustavo Manso
MIT
Sloan School
manso@mit.edu

1 Introduction

Libertarian paternalism, as posed by Thaler and Sunstein (2003, 2008), is arguably one of the most provocative policy contributions in the last two decades. Its beauty stems from its link between two ideas that are on the surface contradictory, but may indeed be an uncompromising compromise. Libertarian paternalism allows a social planner to direct market participants through default options without imposing his will, so that everyone may enjoy the best of both worlds: guidance without the tax of obtrusion.

Not everyone agrees that such a policy is innocuous. For example, Glaeser (2006) argues that libertarian paternalism may also cause bad decisions, is harder to publicly monitor, and may inevitably lead to hard paternalism. Korobkin (2009) argues that, even though libertarian paternalism may induce individuals to make optimal decisions for themselves, collective welfare may decrease. These objections raise an obvious question: When do we expect libertarian paternalism to be welfare improving?

To explore this, we analyze an important dimension of this debate: the effect of libertarian paternalism on information acquisition and social learning. We know from Madrian and Shea (2001) that default options provide information to market participants, which may change both their perceptions and resultant actions. Such intervention also impacts the effectiveness of learning through social interaction (e.g., Duflo and Saez, 2003). So, if learning from others and incentives to acquire information decrease sufficiently when people are guided by a social planner, whether they are forced to make choices or not, total welfare may decrease. This implies that in some circumstances it may be optimal to either implement a limited form of libertarian paternalism or to leave market participants alone, even if some people's choices end up regrettably suboptimal.

We characterize some settings in which providing default options may decrease welfare because information acquisition and aggregation slows. We do this both when information percolates according to a social learning technology (e.g., Duffie and Manso, 2007) and in a market setting in which informed participants can sell their information to others.

In the model that we analyze, there is a continuum of heterogeneous individuals, who each share a common characteristic that is known by the government. The government has to decide whether to disclose this information through a default option or to keep the information to themselves. Individuals may also exert costly effort to find out their own types, which includes the government's information, so that they make an even better decision. As a group, higher aggregate effort also decreases the costs for any one individual to become informed. This form of social learning provides an externality where one individual's effort affects other peoples' welfare and vice versa.

We derive conditions under which default options are optimal and describe when they destroy social surplus. When the information-sharing technology is sufficiently effective, the cost of information acquisition is low, and/or the agent-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 take place. Alternatively, if the information known by the planner is relatively more valuable and these other conditions do not hold, then default options add value.

This sheds light on when libertarian paternalism is likely to add value. Default options are likely to be welfare-improving when individuals are homogeneous. For example, consider the default option of organ donation following a lethal car accident. There is little variation in the quality of healthy organs from different individuals following an accident. In this case, donation as a default is likely to add value. Default options are also likely to be welfare-improving when the information acquired by the planner is relatively valuable compared to the information gathered by individuals. This motivates why default options to participate in a 401(k) retirement plan are so useful. However, default options are unlikely to increase social welfare when peoples' 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 may be portfolio allocation problems. If providing defaults for this decision decreases some peoples' incentives to become savvy, this may lead to a drop in welfare.

We proceed to consider what happens when the government acquires imperfect information about its constituents. In this case, systematic errors decrease the accuracy for people who use the default options, but increase the effort that individual's employ to acquire and aggregate information. We show that the latter effect dominates the former in that issuing no default is more likely to be of value when the government's information is imperfect. Therefore, our analysis addresses the objection raised by Glaeser that social planners are not immune from making errors or having biases.

Given this, we then consider whether the government would ever want to issue an imperfect default even though they have perfect information. We show this not to be the case. That is, despite being given a broader action space including noisy defaults, the government's optimal choice is binary: either issue a fully informative default option or leave individuals to fend for themselves. The same comparative statics still hold as before, which supports the generality of our findings.

Finally, we characterize a market in which information sales are allowed to take place. A fraction of the population are recognized as information gatherers (e.g., brokers in financial markets), whereas the remainder rely on advice markets for guidance. The government faces the same problem as before, and information gatherers decide how much costly effort to employ in accumulating knowledge. The difference here is that there is no social learning technology. Rather, information gatherers may sell their information to the rest of the public for a price. In this version of the model, the presence of a default option decreases the value of advice. That is, since fewer information gatherers will become knowledgeable, the quality of advice in the market suffers. As in the base model, not offering a default option dominates issuing a default option if the cost of effort is low and the value of agent-specific (government) information is high (low). This finding is not dependent on market power, that is, whether the industry is competitive or the advisors have local monopoly power.

The analysis in this paper contributes to the literature on the distortions of paternalism, whether hard or soft (i.e., libertarian paternalism).[1] Bentham (1781) and Hayek (1945) argue that despite the best intentions of a social planner, individuals have both an advantage in gathering precise information about themselves and a greater incentive to do so. As such, any policy that impedes such progress is welfare decreasing. More recently, hard paternalism has been studied by O'Donoghue and Rabin (2003, 2006) who analyze the unintended consequences of sin taxes. Likewise, Camerer et al. (2003) study a form of asymmetric paternalism to minimize such distortions. As mentioned previously, with the recent advent of soft paternalism (e.g., Thaler and Sunstein, 2003), several critics have been wary of unintended consequences induced by such policies (e.g., Glaeser 2006; Korobkin, 2009; Zanitelli, 2009). Our analysis adds to this literature by showing when and where soft paternalism is likely to work and when it is likely to destroy social surplus. Consistent with Bentham (1781) and Hayek (1945), if the knowledge of the government pales in importance compared to what individuals need to know about themselves, libertarian paternalism is likely to be suboptimal. However, if people are more homogeneous and the government's information is accurate and valuable, employing soft paternalism is optimal.

Our work also builds on the work by Carroll et al. (2008) who study optimal default options in a dynamic model, given that individuals tend to procrastinate in making important decisions. They show that default options can function as a control device: when individuals have a hyperbolic discount function, offering a biased default induces people to opt out and make educated decisions sooner. Our focus in this paper is obviously different. We consider that there is information content

[1]See Amir and Lobel (2009) for a recent review of this literature.

in default options, and that this may effect information acquisition and aggregation by people in the market. We show that the government will either issue a fully informative default or no default at all, but never discloses only a portion of their information.

The remainder of the paper is organized as follows. Section 2 outlines our basic model and determines when it is optimal to use default options. In Section 2.1, we consider that the government can only issue fully informative default options, whereas in Section 2.2 we consider imperfect default options. In Section 3, we allow for information sales. Section 4 provides some concluding remarks.

2 Social Learning

2.1 Basic Model

The economy is composed of a government and a continuum (a non-atomic finite measure space $(I, \mathcal{I}, \gamma)$) of heterogeneous, rational individuals who all face a significant economic 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 type. The type $\tilde{\tau}_i$ is the sum a component $\tilde{g}$ that is common to all individuals and an idiosyncratic component $\tilde{t}_i$ that is specific to individual i. We assume that $\tilde{g}$ and $\tilde{t}_i$ are two independent normally distributed random variables, each with zero mean and respective variances Σ_g and Σ_t, and that $\text{Cov}(\tilde{t}_i, \tilde{t}_j) = \rho\Sigma_t$, with $\rho \in [0, 1)$, for any $\{i, j\} \in I^2$ with $i \neq j$.[2] Thus, for each individual i, $\tilde{\tau}_i$ is normally distributed with a mean of zero and a variance of $\Sigma_\tau \equiv \Sigma_g + \Sigma_t$. 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 the expected loss that they suffer.

Before choosing x_i, each individual i can exert some effort in order to improve the probability that he finds out about his own type. An individual who selects an effort level $e_i \in [0, 1]$ observes his true type $\tilde{\tau}_i$ (i.e., receives an informative signal) with probability e_i, and observes nothing otherwise. Individuals know when they did not receive an informative signal. An individual's effort

[2] Note that the positive correlation across the idiosyncratic component t_i of individuals' types does not play a role until we allow for information sales, in section 3.

of e_i comes with a personal utility cost of

$$C(e_i) = \frac{c}{2}(e_i^2 - \alpha \bar{e}^2), \tag{2}$$

where $\bar{e} \equiv \int_I e_i d\gamma$, $\alpha \in [0, 1)$, and c is a positive constant. Given that $\bar{e}$ represents the average effort exerted by individuals in the population, the cost specification in (2) implies that it is cheaper to learn one's 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 that ultimately make it easier for anyone to learn about the economic decision that they have to make. While not specifically modeled, the micro-foundation for this setup might be a model of search in which the distance that a person travels to gather information decreases as more of the population is informed. The parameter α measures the degree of this information externality.

The government costlessly observes the common component $\tilde{g}$ of the individual's types. For example, this could correspond to the government having an informed opinion about the optimal average savings rate for a group of individuals. The government then chooses whether to set a default option that takes $\tilde{g}$ into account or to leave individuals to their own devices. Its goal in this choice is to maximize total welfare. Since individuals are rational, they are able to glean information about $\tilde{g}$ from a default option if it is offered.[3,4] This, in turn, will affect their choice of effort in gathering further information.

Let $\mathcal{S}_i$ denote the information set of an 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 government sets a default option.[5] When there is a default option and the individual does not observe his type, $\mathcal{S}_i = \{\tilde{g}\}$. Finally, when there is no default option and the individual does not observe his type, $\mathcal{S}_i = \varnothing$. The following lemma defines the optimal choice of x_i, given the information set $\mathcal{S}_i$.

Lemma 1. *The optimal choice of x_i for individual i is* $\mathrm{E}\big[\tilde{\tau}_i \mid \mathcal{S}_i\big]$.

With a default option, each individual i who observes an informative signal opts out of the default and chooses $x_i = \tilde{\tau}_i$, whereas any individual who remains uninformed does not opt out, i.e.,

[3] As long as the government's choice for the default option is one-to-one with $\tilde{g}$, every individual can infer $\tilde{g}$ perfectly. Thus it is without loss of generality that we assume in what follows that the government announces $\tilde{g}$ as the default option when it makes such an option available.

[4] A key difference between our model and that of Carroll et al. (2008) is that each individual i is free to extract the benefit from the information in $\tilde{g}$ and to use it for his decision x_i, without incurring any penalty for not choosing $x_i = \tilde{g}$.

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

chooses $x_i = \tilde{g}$, as prescribed by the government. If no default option is offered by the government, any individual i who becomes informed still chooses $x_i = \tilde{\tau}_i$, and chooses $x_i = 0$ if they do not get to observe an informative signal. Consistent with Madrian and Shea's (2001) empirical findings, there is information content in the default options that the government provides, as uninformed individuals optimally (and rationally) choose to use them.

Before choosing x_i but after the government'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 1. It also depends on the individual i's information set $\mathcal{S}_i^0$ which is then $\tilde{g}$ if the government makes a default option available and is empty otherwise. The following lemma summarizes and simplifies this maximization problem.

Lemma 2. *Individual i chooses his effort level e_i to maximize*

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i) \mid \mathcal{S}_i^0\Big] = -(1-e_i)\Big[(1-\delta)\Sigma_g + \Sigma_t\Big] - \frac{c}{2}(e_i^2 - \alpha\bar{e}^2), \tag{3}$$

where $\delta = 1$ when a default option $\tilde{g}$ is offered by the government and $\delta = 0$ when people are left to their own devices.

This result highlights the tradeoff faced by each individual. Effort is costly (second term in (3)) but it reduces the variance that the individual is subject to (first term in (3)). At the same time, the concerted effort of every individual creates a public good, $\bar{e}$, that lowers costs for everyone. Going forward, we make the following assumption, which guarantees an interior solution to the effort problem but does not affect the economics of the analysis.

Assumption 1. *The cost parameter c is such that $c > \Sigma_g + \Sigma_t$.*

The following proposition characterizes the effort choice of individuals, with and without a default option.

Proposition 1. *If the government adopts a default option, each individual chooses effort*

$$e_i^{\mathrm{D}} = \frac{\Sigma_t}{c}, \tag{4}$$

whereas if the government does not adopt a default option, each individual chooses effort

$$e_i^{\mathrm{N}} = \frac{\Sigma_g + \Sigma_t}{c}. \tag{5}$$

Inspection of (4) and (5) shows that individuals exert more effort with higher Σ_t and lower c. That is, the more variance about an individual's type that is resolved when an informative signal is obtained and the lower the cost of acquisition, the more effort each individual is willing to employ. Importantly, it is also the case that

$$e_i^{\mathrm{N}} = e_i^{\mathrm{D}} + \frac{\Sigma_g}{c}.$$

This implies that people exert more effort without a default option, and that the difference between e_i^{N} and e_i^{D} increases as Σ_g gets larger and as c gets smaller. Since the positive externality $\bar{e}$ comes from the average effort of people in the economy, it follows that there are greater opportunities for people to learn from each other when default options are not provided by the government. 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 government has in its possession.

Given Proposition 1 and Lemma 2, we can compute the total welfare with a default option as

$$W^{\mathrm{D}} = -\Sigma_t + \frac{(1+\alpha)\Sigma_t^2}{2c}, \tag{6}$$

and the total welfare without a default option as

$$W^{\mathrm{N}} = -(\Sigma_g + \Sigma_t) + \frac{(1+\alpha)(\Sigma_g + \Sigma_t)^2}{2c}. \tag{7}$$

The next proposition compares welfare with and without a default option.

Proposition 2. *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 in the following region:*

$$\Sigma_g + \Sigma_t < c < (\Sigma_g + 2\Sigma_t)\,\frac{1+\alpha}{2}. \tag{8}$$

This region is non-empty if and only if

$$\Gamma \equiv \frac{\Sigma_g}{\Sigma_g + \Sigma_t} < \frac{2\alpha}{1+\alpha}. \tag{9}$$

According to Proposition 2, welfare without a default option may be higher than welfare with a default option. This arises because the presence of a default option reduces people's incentives to learn about the economic problem they face, which in turn slows the pace of information propagation throughout the economy. 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 learn a lot from each other (i.e., when α is large), and when the cost of information acquisition is low (i.e., when c is small). As shown in (9), the availability of a default option is more likely to be detrimental if the portion

Γ of the volatility that the government can eliminate with its information about $\tilde{g}$ is small relative to the extent of information externalities.

To gain further insight into this result, let us use (6) and (7) and define the difference

$$\Delta W \equiv W^{\mathrm{N}} - W^{\mathrm{D}} = -\Sigma_g + \frac{(1+\alpha)\Sigma_g(\Sigma_g + 2\Sigma_t)}{2c}. \tag{10}$$

Notice that, since $\Sigma_g = \Gamma\Sigma_\tau$ and $\Sigma_t = (1-\Gamma)\Sigma_\tau$, we can rewrite this expression as

$$\Delta W = -\Gamma\Sigma_\tau + \frac{(1+\alpha)\Gamma(2-\Gamma)\Sigma_\tau^2}{2c}. \tag{11}$$

It is easy to verify that, holding the total variance Σ_τ fixed, we have

$$\frac{\partial(\Delta W)}{\partial\Gamma} = -\Sigma_\tau + \frac{(1+\alpha)(1-\Gamma)\Sigma_\tau^2}{c}, \tag{12}$$

and this quantity is positive if and only if

$$\Gamma < 1 - \frac{c}{(1+\alpha)\Sigma_\tau}.$$

That is, an increase in the ability of the government to curb variance by revealing its knowledge of $\tilde{g}$ through a default option makes this option relatively less appealing when Γ is small or the total variance Σ_τ is large. In other words, when important information about individuals is unobservable to the government (small Γ) or when there is a lot of uncertainty about the individuals' economic decision (large Σ_τ), increasing the precision of this information makes default options less appealing, as such options then have a particularly detrimental effect on information gathering incentives, and in turn on information sharing.

Similarly, after fixing the proportion Γ of the total variance that the government can control, we have

$$\frac{\partial(\Delta W)}{\partial\Sigma_\tau} = -\Gamma + \frac{(1+\alpha)\Gamma(2-\Gamma)\Sigma_\tau}{c}, \tag{13}$$

which is positive if and only if

$$\Sigma_\tau > \frac{c}{(1+\alpha)(2-\Gamma)}.$$

Thus an increase in overall uncertainty renders the presence of default options detrimental to welfare when this uncertainty is large to begin with (large Σ_τ) and when the government's ability to reduce uncertainty is limited (small Γ). The former effect has two potential interpretations. First, Σ_τ might proxy for the amount of heterogeneity in the population: when peoples' needs or attributes differ a lot, default options are more likely to be suboptimal. Second, Σ_τ might also proxy for the economic value at risk in each individual's decision: when decisions are more important, the government should refrain from issuing a default in order to promote learning and information sharing by individuals. The latter effect is directly related to the information gathering incentives of individuals: an increase in Σ_τ makes the default option damaging when Γ is small because the importance of the information that individuals forego by exerting less effort to gather it, $(1-\Gamma)\Sigma_\tau$, is large relative to the precision of the information they learn from the default option, $\Gamma\Sigma_\tau$. Together, these comparative statics suggest venues in which default options are likely to add value. For instance, default options are more likely to add value when there is little cross-sectional variation in the population (e.g., healthy organ donations following lethal accidents) than when this variation is higher (e.g., portfolio allocation problems).

By inspection of (10), the relationship between ΔW and Σ_g is non-monotonic. Similarly, our analysis of (12) shows that ΔW is non-monotonically affected by changes in Γ. Based on this, it is feasible that the government can optimize welfare by limiting its information collection to an imperfect signal and by offering to the population a default option that is not perfectly correlated with $\tilde{g}$. We explore this next.

2.2 Imperfect Government

One of Glaeser's (2006) objections to the optimality of libertarian paternalism is that the government may, like individuals, make errors in judgement and decision making. For example, the government might have limited precision when gathering information about its constituents. In this case, default options reveal an imperfect, yet unbiased, signal about $\tilde{g}$.[6] Alternatively, the government may gather perfect information about $\tilde{g}$, but wish to disclose an imperfect signal of their information through a default option. Characterizing these issues is the purpose of this section. Specifically, we first analyze what happens when the government observes $\tilde{g}$ imperfectly, and then consider the government's incentives to fully disclose $\tilde{g}$ even when it observes its value perfectly.

Suppose that the government only observes a noisy signal $\tilde{s} = \tilde{g} + \tilde{\epsilon}$, where $\tilde{\epsilon}$ is normally distributed with mean zero and variance Σ_ϵ, and is independent from $\tilde{g}$ and $\tilde{t}_i$ for all $i \in I$. As before, each individual i can exert effort e_i for a cost given by (2) and learns his type τ_i with probability e_i. If the government issues a default option that conveys its noisy signal, rational

[6] Such mistakes might also result from a systematic bias in the government's information gathering process. Of course, since agents in our model are fully rational, they would correctly interpret the information contained in default options and remove the effects of these systematic biases. With an unbiased, noisy signal about $\tilde{g}$, improving precision is not possible but does induce individuals to employ more effort in acquiring their own information. Therefore, the model as posed could include a systematic bias, but this would not change the economics of our results. Only if individuals could not understand and adjust for the government's biases would such mistakes change the analysis and lead to lower welfare.

individuals will take this into account when choosing how much information to acquire and share. We characterize this effect in the following proposition.

Proposition 3. *If the government implements an imperfect default option with noise Σ_ϵ, each individual i chooses effort*

$$e_i = \frac{(1-\delta)\Sigma_g + \Sigma_t}{c}, \tag{14}$$

where $\delta \equiv \frac{\Sigma_g}{\Sigma_g + \Sigma_\epsilon}$. An individual i who observes a fully informative signal opts out of the default option and chooses $x_i = \tilde{\tau}_i = \tilde{g} + \tilde{t}_i$. An individual i who does not become informed chooses $x_i = \delta\tilde{s} = \delta\big(\tilde{g} + \tilde{\epsilon}\big)$, the default option offered by the government.

As in Proposition 1, the optimal choice of effort is strictly decreasing in c and increasing in Σ_t and Σ_g. Additionally, as the amount of noise in the default increases (higher Σ_ϵ, and thus lower δ), the higher is the effort that each individual is willing to exert to learn about $\tilde{\tau}_i$. Therefore, the precision of information contained in the default option drives the incentives of individuals to acquire information, which in turn affects how much is learned via information sharing.

Given Proposition 3, we can compute the total welfare with a noisy default option as

$$\begin{aligned} W^{\mathrm{D}}(\Sigma_\epsilon) &= -\big[(1-\delta)\Sigma_g + \Sigma_t\big] + \frac{\big[(1-\delta)\Sigma_g + \Sigma_t\big]^2}{2c}(1+\alpha) \\ &= -\left(\frac{\Sigma_\epsilon\Sigma_g}{\Sigma_g + \Sigma_\epsilon} + \Sigma_t\right) + \frac{\left(\frac{\Sigma_\epsilon\Sigma_g}{\Sigma_g+\Sigma_\epsilon} + \Sigma_t\right)^2}{2c}(1+\alpha). \end{aligned} \tag{15}$$

The next proposition compares welfare with and without a default option when the government's information is imperfect.

Proposition 4. *The total welfare W^{N} without a default option is higher than the total welfare $W^{\mathrm{D}}(\Sigma_\epsilon)$ with a noisy default option if the cost parameter c is in the following region:*

$$\Sigma_g + \Sigma_t < c < \left(\frac{2\Sigma_\epsilon + \Sigma_g}{\Sigma_\epsilon + \Sigma_g}\Sigma_g + 2\Sigma_t\right)\frac{1+\alpha}{2}. \tag{16}$$

This region is non-empty if and only if

$$\Gamma\Phi < \frac{2\alpha}{1+\alpha}, \tag{17}$$

where $\Gamma = \frac{\Sigma_g}{\Sigma_g+\Sigma_t}$ and $\Phi = \frac{\Sigma_g}{\Sigma_g+\Sigma_\epsilon}$.

Comparing the result in Proposition 4 with that in Proposition 2, the region in which no default dominates default is larger when the government's information is imprecise. In fact, Proposition 4 shows that this region gets larger as Σ_ϵ increases (and Φ decreases). This result is not obvious: an imprecise default hurts individuals who decide take the default option, but also provides incentives for individuals to search more intensively, which improves information sharing. Comparison of Propositions 2 and 4 shows that the first effect dominates the second, confirming Glaeser's (2006) conjecture that the case for libertarian paternalism is weaker if the government makes errors in judgement or has imprecise information.

Clearly, this motivates an analysis of whether the government would optimally choose to issue a noisy default, even when they have (or have free access to) perfect information about $\tilde{g}$. Thus let us consider a broader action space for the government in which it can issue default options that do not convey a precise signal regarding $\tilde{g}$. As such, the government could still choose to issue a default option that conveys $\tilde{g}$ perfectly, but we now allow it to instead issue a default option that conveys $\tilde{g} + \tilde{\epsilon}$, in which the government chooses the variance $\Sigma_\epsilon > 0$ of $\tilde{\epsilon}$. If a finite Σ_ϵ is chosen, individuals can learn some (i.e., incomplete) information about their decision from the default. Of course, as before, the government can still make the default option perfectly informative about $\tilde{g}$ by choosing $\Sigma_\epsilon = 0$, and effectively refrain from making a default option available by choosing $\Sigma_\epsilon = \infty$.

Given our previous discussion, the government's choice of Σ_ϵ affects welfare through two channels. A higher precision improves the choices that individuals make when they do not observe an informative signal, but it decreases the incentives of individuals to collect and share information in the first place. Taking these two forces into account, the next proposition characterizes the government's optimal default policy.

Proposition 5. *The optimal choice of noisy default policy is given by*

$$\Sigma_\epsilon^* = \begin{cases} 0, & \text{if } c > (\Sigma_g + 2\Sigma_t)\frac{1+\alpha}{2} \\ \infty, & \text{otherwise.} \end{cases} \tag{18}$$

Proposition 5 implies that our analysis in Section 2.1 holds even when we consider a broader action space for the government. That is, the government's decision is effectively binary: it either chooses a fully informative default option or it offers no default whatsoever. Again, if the cost of information acquisition is sufficiently high (high c), the size of the variation or value at risk is sufficiently low (low $\Sigma_g + \Sigma_t$), or the information sharing technology is sufficiently weak (low α), the government issues a fully informative default option (i.e., $\Sigma_\epsilon = 0$). Otherwise, it lets individuals fend for themselves (i.e., $\Sigma_\epsilon = \infty$).

3 Information Sales

So far, our model shows that the adoption of default options is costly and potentially suboptimal when individuals in the economy can help each other learn about their own type. In this section, we show that the externality need not be of the form specified in section 2. In particular, we show that allowing individuals to sell their information to uninformed individuals can generate similar results. That is, the presence of default options reduces the incentive for individuals to gather and resell their information. This leads to an overall reduction of information 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) purchase information from other individuals in the economy. More specifically, we assume that a subset $I_\mu \in I$, with $\gamma(I_\mu) = \mu$, of individuals are skilled in the sense that they can gather information about their type with the same technology as before; that is, for a cost of $C(e_i)$, individual $i \in I_\mu$ receives a signal that reveals his $\tilde{g} + \tilde{t}_i$ with probability e_i. To emphasize the fact that externalities derive purely from information sales, we set $\alpha = 0$ in (2), so that $C(e_i) = \frac{c}{2}e_i^2$. The other individuals, $j \in I \setminus I_\mu$, are unskilled in the sense that gathering information about their own type is prohibitively costly.

Instead, these unskilled individuals can purchase information from skilled individuals. 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 $\tilde{g} + \tilde{t}_i$ 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 $\tilde{g} + \tilde{t}_i$ (which is correlated with his own type $\tilde{g} + \tilde{t}_j$) or noise (which is not).[7] Throughout this section, we go back to the assumption that the government's default option is perfect (i.e., equal to $\tilde{g}$) when it is made available; that is, we refrain from showing as in section 2.2 that this choice is optimal even if the government can choose the precision of its information. The following lemma characterizes the value derived from the information by an unskilled individual who consults a randomly selected skilled individual.

Lemma 3. *If the government 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*

[7]We assume that skilled individuals who do not learn their own type sell a randomly drawn number from a normal distribution with a mean of zero and a variance of $\Sigma_g + \Sigma_t$, which makes it impossible for information buyers to tell noise from real information.

is

$$\nu_0 = \frac{(\Sigma_g + \rho\Sigma_t)^2}{\Sigma_g + \Sigma_t}\bar{e}_\mu = \left[\Gamma + \rho(1-\Gamma)\right]^2\Sigma_\tau\bar{e}_\mu, \tag{19}$$

where $\bar{e}_\mu \equiv \frac{1}{\mu}\int_{I_\mu} e_i d\gamma$, $\Sigma_\tau = \Sigma_g + \Sigma_t$, *and* $\Gamma = \frac{\Sigma_g}{\Sigma_g+\Sigma_t}$. *If the government 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*

$$\nu_1 = \rho^2\Sigma_t\bar{e}_\mu = \rho^2(1-\Gamma)\Sigma_\tau\bar{e}. \tag{20}$$

Unskilled individuals are willing to pay more to learn a skilled individual's type when they know that skilled individuals exert a lot of effort to learn their own type, i.e., ν_0 and ν_1 are both increasing in $\bar{e}$. This makes sense as a fraction $\bar{e}_\mu$ of the μ skilled individuals will be informed in equilibrium, while the other $1 - \bar{e}_\mu$ skilled individuals sell useless noise. From (19) and (20), we can also see that unskilled individuals are willing to pay a higher price for a skilled individual's information when their type is highly variable (large Σ_τ) and when it is more highly correlated with that of other individuals (large ρ). This last result is consistent with the fact that, keeping Σ_τ fixed, ν_0 is increasing in Γ, as types are more correlated when the common component $\tilde{g}$ accounts for a larger portion of each individual's type. This is also consistent with ν_1 being decreasing in Γ as, when the government announces $\tilde{g}$, the unknown portion of an individual's type correlates with someone else's type only to the extent that the default option leaves uncertainty regarding $\tilde{t}_i$. In fact, using (19) and (20), it is straightforward to verify that $\nu_0 > \nu_1$ for a given total variance Σ_τ and aggregate level of effort $\bar{e}_\mu$. Indeed, because types are more correlated across individuals when $\tilde{g}$ 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 by the government. As we shall see below, this difference between ν_0 and ν_1 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 traders to consult another skilled individual. To capture these possibilities in a tractable manner, we assume that each unskilled individual meets with one randomly selected skilled individual, and that the economic surplus from their transaction is split as a Nash bargaining outcome. More specifically, we assume that a skilled individual charges $p = \theta\nu_\delta$, where $\theta \in [0,1]$ and $\delta = 1$ if a default option is made available, for the information he sells to an unskilled individual. When $\theta = 1$ ($\theta = 0$), the skilled (unskilled) individual extracts all the surplus from the transaction.

Setting $\theta \in (0,1)$ allows us to capture any intermediate market power scenario. As the following analysis shows, our results are unaffected by the size of θ, as money exchanges between individuals cancel out in the total welfare function that the government seeks to maximize. We start with the following result, which describes the equilibrium in the absence of a default option.

Proposition 6. *If the government does not adopt a default option, then each skilled individual* $i \in I_\mu$ *chooses an effort level* $e_i = \frac{\Sigma_g+\Sigma_t}{c} = \frac{\Sigma_\tau}{c}$, *and chooses* $x_i = \tilde{\tau}_i$ *or* $x_i = 0$, *depending on whether or not he observes* $\tilde{\tau}_i$. *Each unskilled individual* $j \in I \setminus I_\mu$ *purchases a signal* $\tilde{s}_j = \tilde{\tau}_{\tilde{\imath}}$ *from a randomly selected skilled individual* $\tilde{\imath} \in I_\mu$ *for a price* $p = \theta\nu_0$, *with* ν_0 *given by (19), and chooses*

$$x_j = \frac{\Sigma_g + \rho\Sigma_t}{\Sigma_g + \Sigma_t}\bar{e}_\mu \tilde{s}_j = \big[\Gamma + \rho(1-\Gamma)\big]\bar{e}_\mu \tilde{s}_j. \tag{21}$$

The skilled individuals' behavior is the same as in section 2.1. In particular, their behavior is not affected by the possibility of reselling their information to unskilled agents. This is due to the fact that unskilled agents cannot distinguish between skilled individuals who learn their type and skilled individuals who do not. That is, they pay $\theta\nu_0$ to the one skilled individual they encounter, informed or not. As we see from (21), 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 $\bar{e}_\mu$ all ultimately lead to a higher correlation between $\tilde{s}_j$ and $\tilde{\tau}_j$. The following result is the analogue of Proposition 6 when the government makes a default option $\tilde{g}$ available.

Proposition 7. *If the government adopts a default option, then each skilled individual* $i \in I_\mu$ *chooses an effort level* $e_i = \frac{\Sigma_t}{c} = \frac{(1-\Gamma)\Sigma_\tau}{c}$, *and chooses* $x_i = \tilde{\tau}_i$ *or* $x_i = \tilde{g}$, *depending on whether or not he observes* $\tilde{\tau}_i$. *Each unskilled individual* $j \in I \setminus I_\mu$ *purchases a signal* $\tilde{s}_j = \tilde{\tau}_{\tilde{\imath}}$ *from a randomly selected skilled individual* $\tilde{\imath} \in I_\mu$ *for a price* $p = \theta\nu_1$, *with* ν_1 *given by (20), and chooses*

$$x_j = \tilde{g} + \rho\bar{e}_\mu(\tilde{s}_j - \tilde{g}). \tag{22}$$

The comparative statics on the individuals' choices with respect to Σ_τ, ρ and $\bar{e}_\mu$ are similar to those in Proposition 6: more risk (large Σ_τ) leads to more effort, and more correlation (large ρ and $\bar{e}_\mu$) 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 information price (through ν_1, as shown in Lemma 3), 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 component included in the signal and only place weight on the idiosyncratic component. 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 ν_0 and ν_1. 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 government's default option, we compare total welfare with and without this option. In this case, welfare must be aggregated between skilled and unskilled individuals. This is done in the following lemma.

Lemma 4. *The total welfare without a default option is*

$$W^{\mathrm{N}} = -(\Sigma_g + \Sigma_t) + \frac{\mu}{2c}(\Sigma_g + \Sigma_t)^2 + \frac{1-\mu}{c}(\Sigma_g + \rho\Sigma_t)^2. \tag{23}$$

The total welfare with a default option is

$$W^{\mathrm{D}} = -\Sigma_t + \frac{\mu}{2c}\Sigma_t^2 + \frac{1-\mu}{c}\rho^2\Sigma_t^2. \tag{24}$$

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 (23) and (24) 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

$$\frac{\partial W^{\mathrm{N}}}{\partial \rho} = \frac{2(1-\mu)}{c}(\Sigma_g + \rho\Sigma_t)\Sigma_t > 0 \tag{25}$$

and

$$\frac{\partial W^{\mathrm{D}}}{\partial \rho} = \frac{2(1-\mu)}{c}\rho\Sigma_t^2 > 0. \tag{26}$$

That is, a larger correlation across individuals' types leads to more welfare when an 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., $1-\mu$ is large). Finally, it is clear that (25) is greater than (26): the advice channel is more crucial, and the role of ρ greater, when the government refrains from making a default option available, as unskilled individuals can then rely only on the skilled agents' information for their decisions.

The next proposition, which is our last result, is the analogue of Proposition 2 when we allow for information sales.

Proposition 8. *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 in the following region:*

$$\Sigma_g + \Sigma_t < c < \left(1 - \frac{\mu}{2}\right)\Sigma_g + \left[\mu + 2(1-\mu)\rho\right]\Sigma_t. \tag{27}$$

This region is non-empty if and only if $\rho > \frac{1}{2}$ *and*

$$\frac{\Sigma_g}{\Sigma_t} < \frac{2(1-\mu)}{\mu}(2\rho - 1). \tag{28}$$

As mentioned above, the role of ρ plays an especially important welfare role in information sales when the government does not make a default option available. Proposition 8 formalizes this by showing that $\rho \leq \frac{1}{2}$ always makes the availability of a default option optimal. That is, unskilled individuals are better off learning the common component of their type perfectly from the government when the information that can be acquired from other agents 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.

We can also see from (28) that default options are less valuable when Σ_t is large and Σ_g is small, which is similar to our findings in section 2. The extent to which the government can resolve the uncertainty faced by the population is still an important determinant of the usefulness of default option. Interestingly, however, default options are more valuable when a larger fraction of the population is skilled (large μ), even when ρ is large. This arises because the information externalities that skilled individuals bring to the economy 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 agents with and without a default option (as derived in Proposition 7) do not lead to significantly different externalities.[8]

4 Concluding Remarks

Libertarian paternalism is an alluring idea because it combines two policies that appear incompatible at first glance, but work well together in many economic settings. However, one needs to be cautious when implementing the ideals of such a policy in practice. As we show in our analysis, it is not necessarily the paternalistic partner in this union that causes problems in the relationship.

[8]Note that an effort cost function that includes, as in (2), an externality component in the information acquisition process of the skilled individuals would mitigate this result. Indeed, $\alpha > 0$ would lead to both skilled and unskilled benefitting from a more concerted information acquisition effort.

Rather, the freedom that participants exercise in the market may lead to side effects that decrease social welfare.

Indeed, as its name suggests, libertarian paternalism preserves the rights of individuals to act in their own best interest, free-ride on 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 Madrian and Shea (2001) demonstrate, default options have information content, which participants may take into consideration when making key decisions. Importantly, this may affect incentives to gather further information, which in turn may alter the success of information aggregation, either through social learning or information sales in the market.

We explore when libertarian paternalism is more or less likely to add value given this externality. We show that default options are more likely 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 government knows less about its constituents, when people are heterogeneous, and when the value at stake in the decision is large.

Our analysis adds to previous work by Carroll et al. (2008) and increases our understanding of default options and the 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

Proof of Lemma 1

Individual i must choose x_i in order to maximize

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

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

Proof of Lemma 2

Let $\delta = 1$ if the government announces a default option $\tilde{g}$ and $\delta = 0$ otherwise. Using Lemma 1, individual i's expected utility is given by

$$\begin{aligned}
\mathrm{E}\Big[\tilde{U}_i(x_i) \mid \mathcal{S}_i^0\Big] &= \mathrm{E}\Big[-(\tilde{\tau}_i - x_i)^2 \mid \mathcal{S}_i^0\Big] = \mathrm{E}\Big\{\mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \mid \mathcal{S}_i\big] \mid \mathcal{S}_i^0\Big\} \\
&= \Pr\big\{\mathcal{S}_i = \{\tilde{\tau}_i\} \mid \mathcal{S}_i^0\big\}\, \mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \mid \tilde{\tau}_i\big] + \Pr\big\{\mathcal{S}_i = \{\tilde{g}\} \mid \mathcal{S}_i^0\big\}\, \mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \mid \tilde{g}\big] \\
&\qquad + \Pr\big\{\mathcal{S}_i = \varnothing \mid \mathcal{S}_i^0\big\}\, \mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2\big] \\
&= e_i\mathrm{E}\big[-(\tilde{\tau}_i - \tilde{\tau}_i)^2\big] + (1 - e_i)\delta\mathrm{E}\big[-(\tilde{\tau}_i - \tilde{g})^2\big] + (1 - e_i)(1 - \delta)\mathrm{E}\big[-(\tilde{\tau}_i - 0)^2\big] \\
&= -(1 - e_i)\delta\Sigma_t + -(1 - e_i)(1 - \delta)(\Sigma_g + \Sigma_t) \\
&= -(1 - e_i)\Big[(1 - \delta)\Sigma_g + \Sigma_t\Big].
\end{aligned}$$

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

Proof of Proposition 1

As shown in Lemma 2, each individual i chooses e_i to maximize

$$-(1 - e_i)\Big[(1 - \delta)\Sigma_g + \Sigma_t\Big] - \frac{c}{2}(e_i^2 - \alpha\bar{e}^2),$$

where $\delta = 1$ when a default option $\tilde{g}$ is offered by the government and $\delta = 0$ otherwise. The first-order condition for this problem is

$$(1 - \delta)\Sigma_g + \Sigma_t - ce_i = 0,$$

which implies that

$$e_i = \frac{(1 - \delta)\Sigma_g + \Sigma_t}{c}.$$

It is easy to see that the second order condition is satisfied and thus the above e_i corresponds to a maximum. The effort levels with and without a default option, (4) and (5), are obtained by setting δ equal to one and zero respectively. ■

Proof of Proposition 2

A simple comparison of (6) and (7) yields the second inequality in (8). The first inequality in (8) comes from Assumption 1. The region is non-empty if and only if

$$\Sigma_g + \Sigma_t < (\Sigma_g + 2\Sigma_t)\,\frac{1+\alpha}{2},$$

which simplifies to the condition in (9). ■

Proof of Proposition 3

Using the projection theorem for normal variables, it is straightforward to show that $\mathrm{E}\big[\tilde{\tau}_i \,|\, \tilde{s}\big] = \frac{\Sigma_g}{\Sigma_g+\Sigma_\epsilon}\tilde{s} = \delta\tilde{s}$ and $\mathrm{Var}\big[\tilde{\tau}_i \,|\, \tilde{s}\big] = \Big(1 - \frac{\Sigma_g}{\Sigma_g+\Sigma_\epsilon}\Big)\Sigma_g + \Sigma_t = (1-\delta)\Sigma_g + \Sigma_t$, where $\delta = \frac{\Sigma_g}{\Sigma_g+\Sigma_\epsilon}$. Thus, when individual i's information set is $\mathcal{S}_i = \{\tilde{s}\}$ at the time of his decision about x_i, Lemma 1 implies that $x_i = \delta\tilde{s}$. When individual i observes his type and $\mathcal{S}_i = \{\tilde{\tau}_i\}$, then he chooses $x_i = \tilde{\tau}_i$, as before. At the time of his effort decision, individual i's information set is $\mathcal{S}_i^0 = \{\tilde{s}\}$, and thus

$$\begin{aligned}
\mathrm{E}\Big[\tilde{U}_i(x_i) \,|\, \mathcal{S}_i^0\Big] &= \mathrm{E}\Big[-(\tilde{\tau}_i - x_i)^2 \,|\, \mathcal{S}_i^0\Big] = \mathrm{E}\Big\{\mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \,|\, \mathcal{S}_i\big] \,|\, \mathcal{S}_i^0\Big\} \\
&= \Pr\big\{\mathcal{S}_i = \{\tilde{\tau}_i\} \,|\, \mathcal{S}_i^0\big\}\,\mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \,|\, \tilde{\tau}_i\big] + \Pr\big\{\mathcal{S}_i = \{\tilde{s}\} \,|\, \mathcal{S}_i^0\big\}\,\mathrm{E}\big[-(\tilde{\tau}_i - x_i)^2 \,|\, \tilde{s}\big] \\
&= e_i\mathrm{E}\big[-(\tilde{\tau}_i - \tilde{\tau}_i)^2\big] + (1-e_i)\mathrm{E}\big[-(\tilde{\tau}_i - \delta\tilde{s})^2 \,|\, \tilde{s}\big] \\
&= -(1-e_i)\mathrm{Var}\big[\tilde{\tau}_i \,|\, \tilde{s}\big] = -(1-e_i)\Big[(1-\delta)\Sigma_g + \Sigma_t\Big].
\end{aligned}$$

Therefore, each individual i chooses e_i to maximize

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i) \,|\, \mathcal{S}_i^0\Big] = -(1-e_i)\Big[(1-\delta)\Sigma_g + \Sigma_t\Big] - \frac{c}{2}(e_i^2 - \alpha\bar{e}^2).$$

The first-order condition for this problem is

$$(1-\delta)\Sigma_g + \Sigma_t - ce_i = 0,$$

which leads to (14). It is easy to verify that the second-order condition is satisfied. ■

Proof of Proposition 4

The first inequality in (16) comes from Assumption 1. Let us define $\Delta W(\Sigma_\epsilon) \equiv W^{\mathrm{N}} - W^{\mathrm{D}}(\Sigma_\epsilon)$. Using (7) and (15), it is easy to show that

$$\Delta W(\Sigma_\epsilon) = \frac{\Sigma_g^2 \Big\{(1+\alpha)\big[\Sigma_g(2\Sigma_\epsilon + \Sigma_g) + 2\Sigma_t(\Sigma_\epsilon + \Sigma_g)\big] - 2c(\Sigma_\epsilon + \Sigma_g)\Big\}}{2c(\Sigma_\epsilon + \Sigma_g)^2}.$$

This quantity is positive if and only if the second inequality in (16) is satisfied. For the region in (16) to be non-empty, we must have

$$\Sigma_g + \Sigma_t < \left(\frac{2\Sigma_\epsilon + \Sigma_g}{\Sigma_\epsilon + \Sigma_g}\Sigma_g + 2\Sigma_t\right)\frac{1+\alpha}{2},$$

which produces condition (17). ■

Proof of Proposition 5

As shown in (15), welfare with a noisy default policy is given by

$$W^{\mathrm{D}}(\Sigma_\epsilon) = -\left(\frac{\Sigma_\epsilon\Sigma_g}{\Sigma_g + \Sigma_\epsilon} + \Sigma_t\right) + \frac{\left(\frac{\Sigma_\epsilon\Sigma_g}{\Sigma_g + \Sigma_\epsilon} + \Sigma_t\right)^2}{2c}(1+\alpha).$$

After taking the derivative of this expression with respect to Σ_ϵ and simplifying, we find

$$\frac{\partial W^{\mathrm{D}}(\Sigma_\epsilon)}{\partial \Sigma_\epsilon} = \frac{\Sigma_g^2}{c(\Sigma_g + \Sigma_\epsilon)^3}\Big\{(1+\alpha)\big[\Sigma_g\Sigma_t + \Sigma_\epsilon(\Sigma_g + \Sigma_t)\big] - c(\Sigma_\epsilon + \Sigma_g)\Big\}. \tag{29}$$

If $c > (1+\alpha)(\Sigma_g + \Sigma_t)$, this derivative is always negative and it is optimal to set Σ_ϵ as low as possible, that is, $\Sigma_\epsilon^* = 0$. If $c < (1+\alpha)\Sigma_t$, the above derivative is always positive and it is therefore optimal to set Σ_ϵ as high as possible, that is $\Sigma_\epsilon^* = \infty$, which is equivalent to the government not offering a default option. Finally, if $(1+\alpha)\Sigma_t < c < (1+\alpha)(\Sigma_g + \Sigma_t)$, then (29) is greater than zero when

$$\Sigma_\epsilon > \frac{c - (1+\alpha)\Sigma_t}{(1+\alpha)(\Sigma_g + \Sigma_t) - c}\Sigma_g,$$

and smaller than zero otherwise. This means that the maximum can only be achieved at $\Sigma_\epsilon = 0$ (i.e., default option without noise) or $\Sigma_\epsilon = \infty$ (i.e., no default option). The optimal default choice must therefore be the same as in Proposition 2, leading to (18). ■

Proof of Lemma 3

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 government does not make a default option available. If $\tilde{s}_j = \tilde{g} + \tilde{t}_i$, then the reduction in variance experienced by individual j from knowing $\tilde{s}_j$ is given by

$$\mathrm{Var}\big(\tilde{g}+\tilde{t}_j\big) - \mathrm{Var}\big(\tilde{g}+\tilde{t}_j \mid \tilde{g}+\tilde{t}_i\big) = (\Sigma_g+\Sigma_t) - \left[\Sigma_g+\Sigma_t - \frac{(\Sigma_g+\rho\Sigma_t)^2}{\Sigma_g+\Sigma_t}\right] = \frac{(\Sigma_g+\rho\Sigma_t)^2}{\Sigma_g+\Sigma_t},$$

where we use the projection theorem to calculate the expression in square brackets. If $\tilde{s}_j$ is pure noise, then individual j does not experience a reduction in variance. Since a fraction $\bar{e}_\mu$ of the skilled traders learn their type, the unconditional reduction in variance experienced by individual j from learning individual i's information is $\frac{(\Sigma_g+\rho\Sigma_t)^2}{\Sigma_g+\Sigma_t}\bar{e}_\mu$, which can be rewritten as $\big[\Gamma+\rho(1-\Gamma)\big]^2\Sigma_\tau\bar{e}_\mu$ using the fact that $\Sigma_g = \Gamma\Sigma_\tau$ and $\Sigma_t = (1-\Gamma)\Sigma_\tau$. Since this quantity represents the increase in expected utility enjoyed by individual j as a result of knowing $\tilde{s}_j$, this is the maximum price that he is willing to pay for it. The case in which the government makes a default option available is similarly derived. ■

Proof of Proposition 6

Let $\tilde{\pi}_i$ denote the profits that a skilled individual $i \in I_\mu$ generates from selling information to unskilled individuals. With an information price $p = \theta\nu_0$, the $1-\mu$ unskilled individuals will pay a total sum of $(1-\mu)p = (1-\mu)\theta\nu_0$ to acquire signals from the μ skilled agents. Since these skilled agents are randomly selected, the expected profits from information sales of any one skilled individual i are

$$\mathrm{E}\big[\tilde{\pi}_i\big] = \frac{(1-\mu)\theta\nu_0}{\mu}.$$

Thus, using the same notation and reasoning as in Lemma 2, this skilled individual i must choose e_i in order to maximize

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i) + \tilde{\pi}_i\Big] = -(1-e_i)(\Sigma_g+\Sigma_t) - \frac{c}{2}e_i^2 + \frac{(1-\mu)\theta\nu_0}{\mu}.$$

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 $e_i = \frac{\Sigma_g+\Sigma_t}{c}$. After purchasing $\tilde{s}_j$ from a skilled agent, unskilled individual j must choose x_j in order to maximize $\mathrm{E}\big[-(\tilde{g}+\tilde{t}_j-x_j)^2 \,|\, \tilde{s}_j\big]$. By Lemma 1, this individual chooses

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

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

Proof of Proposition 7

Let $\tilde{\pi}_i$ denote the profits that a skilled individual $i \in I_\mu$ generates from selling information to unskilled individuals. With an information price $p = \theta\nu_1$, the $1-\mu$ unskilled individuals will pay a total sum of $(1-\mu)p = (1-\mu)\theta\nu_1$ to acquire signals from the μ skilled agents. Since these skilled agents are randomly selected, the expected profits from information sales of any one skilled individual i are

$$\mathrm{E}\big[\tilde{\pi}_i\big] = \frac{(1-\mu)\theta\nu_1}{\mu}.$$

Thus, using the same notation and reasoning as in Lemma 2, this skilled individual i must choose e_i in order to maximize

$$\mathrm{E}\Big[\tilde{U}_i(x_i) - C(e_i) + \tilde{\pi}_i \;\Big|\; \tilde{g}\Big] = -(1-e_i)\Sigma_t - \frac{c}{2}e_i^2 + \frac{(1-\mu)\theta\nu_1}{\mu}.$$

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 $e_i = \frac{\Sigma_t}{c}$. After purchasing $\tilde{s}_j$ from a skilled agent, unskilled individual j must choose x_j in order to maximize $\mathrm{E}\big[-(\tilde{g} + \tilde{t}_j - x_j)^2 \mid \tilde{g}, \tilde{s}_j\big]$. By Lemma 1, this individual chooses

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

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

Proof of Lemma 4

Suppose first that there is no default option. From the proof of Proposition 6, we know that the welfare of any one skilled individual $i \in I_\mu$ is given by

$$W_i^{\mathrm{N}} = -(1-e_i)(\Sigma_g + \Sigma_t) - \frac{c}{2}e_i^2 + \frac{(1-\mu)p}{\mu}.$$

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

$$W_i^{\mathrm{N}} = -(\Sigma_g + \Sigma_t) + \nu_0 - p,$$

and so total welfare is

$$W^{\mathrm{N}} \equiv \int_I W_i^{\mathrm{N}} d\gamma = \int_{I_\mu} \Big[-(1-e_i)(\Sigma_g+\Sigma_t) - \frac{c}{2}e_i^2\Big] d\gamma + \int_{I\setminus I_\mu} \Big[-(\Sigma_g+\Sigma_t)+\nu_0\Big] d\gamma$$
$$= -(\Sigma_g+\Sigma_t) + \int_{I_\mu} \Big[e_i(\Sigma_g+\Sigma_t) - \frac{c}{2}e_i^2\Big] d\gamma + (1-\mu)\nu_0.$$

In equilibrium, we know from Proposition 6 that $e_i = \bar{e}_\mu = \frac{\Sigma_g+\Sigma_t}{c}$, $p = \theta\nu_0$, and $\nu_0 = \frac{(\Sigma_g+\rho\Sigma_t)^2}{\Sigma_g+\Sigma_t}\bar{e}_\mu$. After using these expressions in the total welfare function above, we get

$$W^{\mathrm{N}} = -(\Sigma_g+\Sigma_t) + \mu\Big[\bar{e}_\mu(\Sigma_g+\Sigma_t) - \frac{c}{2}\bar{e}_\mu^2\Big] + (1-\mu)\frac{(\Sigma_g+\rho\Sigma_t)^2}{\Sigma_g+\Sigma_t}\bar{e}_\mu,$$

which simplifies to (23). The calculations are similar with the default option. ■

Proof of Proposition 8

A simple comparison of (23) and (24) yields the second inequality in (27). The first inequality in (27) comes from Assumption 1. The region is non-empty if and only if

$$\Sigma_g+\Sigma_t < \left(1-\frac{\mu}{2}\right)\Sigma_g + \big[\mu + 2(1-\mu)\rho\big]\Sigma_t,$$

which simplifies to the condition in (28). ■

References

Amir, On, and Orly Lobel. 2009. "Stumble, Predict, Nudge: How Behavioral Economics Informs Law and Policy." University of San Diego School of Law, Research Paper 09-006.

Bentham, Jeremy. 1781. *An Introduction to the Principles of Morals and Legislation*. Reprinted by Batoche Books, Kitchener, 2000.

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.

Carroll, Gabriel, James Choi, David Laibson, Brigitte Madrian, and Andrew Metrick. 2008. "Optimal Defaults and Active Decisions." Forthcoming, *Quarterly Journal of Economics*.

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

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.

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

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

Korobkin, Russell. 2009. "Libertarian Welfarism." University of California at Los Angeles, Working Paper.

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.

O'Donoghue, Ted, and Matthew Rabin. 2003. "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes." *American Economic Review*, 93: 186-191.

O'Donoghue, Ted, and Matthew Rabin. 2006. "Optimal Sin Taxes." Forthcoming, *Journal of Political Economy*.

Rizzo, Mario and Douglas Glen Whitman. 2008. "The Knowledge Problem of New Paternalism." New York University, Law and Economics Research Paper Series, Working Paper Number 08-60.

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.

Zanitelli, Leandro Martins. 2009. "Default Rules and the Inevitability of Paternalism." Centro Universitarío Ritter dos Reis, Working Paper.