ESG RECHERCHE

DOCUMENT DE TRAVAIL / WORKING PAPER

No. 2022-06

The Optimal Design of Pensions and Financial Education, when Individuals are Biased and Naive

Chiara Canta et Marie-Louise Leroux

Août 2022

ESG UQÀM
Département des sciences économiques
École des sciences de la gestion
Université du Québec à Montréal

The Optimal Design of Pensions and Financial Education, when Individuals are Biased and Naive

Chiara Canta, TBS Business School

Marie-Louise Leroux, ESG-UQAM, CORE, CESifo and CIRANO

Document de travail No. 2022-06

Août 2022

Département des Sciences Économiques
Université du Québec à Montréal
Case postale 8888,
Succ. Centre-Ville
Montréal, (Québec), H3C 3P8, Canada
Courriel : frechette.karine@uqam.ca
Site web : http://economie.esg.uqam.ca

The optimal design of pensions and financial education, when individuals are biased and naive *

Chiara Canta[†], Marie-Louise Leroux[‡]

July 8, 2022

Abstract

This paper studies the optimal design of a pension system together with publicly-provided individualized financial education. Agents can invest in both a risky and a non risky asset and can either under- or over-estimate the expected return of the risky asset. We show that, under perfect information on the misperception biases, it is optimal for the government to impose a uniform level of pension contributions equal to the first-best level of investment in the non-risky asset and a U-shaped level of education. Under asymmetric information, we show that the level of education is always distorted upward for agents with important misperception biases (who either under- or over-estimate financial returns) but, can be distorted upward or downward for agents with mild misperception biases. Whether we end up in one or the other situation depends on the size of the public and private costs of education as well as on the shape of the distribution of the misperception biases in the economy.

Keywords: Pension systems, financial education, misperception, taxation, asymmetric information.

JEL Codes: D14; H21; H53; H55; I22.

*Financial support from Social Sciences and Humanities Research Council (Grant 435-2020-0787) and by Fonds de Recherche du Québec-Société et Culture (Grant 2019-SE2-252890) is gratefully acknowledged.

[†]TBS Business School. E-mail: c.canta@tbs-education.fr

[‡]Département des Sciences Economiques, ESG-UQAM; CORE (Université catholique de Louvain); CESifo (Munich) and CIRANO (Montreal). E-mail: leroux.marie-louise@uqam.ca

1 Introduction

Despite the historical higher performance of stocks over bonds and banking accounts, agents are still reluctant to hold stock market assets. This is often referred as the "stock market participation puzzle". Economists have long tried to explain that puzzle, invoking, for instance, high participation costs, ambiguity aversion, loss aversion, lack of confidence in the financial system.[1] One explanation may also be that agents misperceive expected stock returns and risk levels. For instance, Kezdi and Willis (2009) show, using HRS data, that stock holding is low because agents expect a lower mean return and a higher variance than the true, historical ones. In the same way, Hurd and Rohwedder (2011) show that agents are in general more pessimistic about stock market returns than historical outcomes justify, and would therefore prefer (riskless) saving accounts. Another reason for explaining that puzzle relates to the low level of financial education of agents. For instance, Van Rooj et al. (2011), in the case of the Netherlands, and Arrondel et al. (2015), in the case of France, show that agents with low financial literacy are less likely to invest in stocks.

Misperception biases regarding the return of financial assets may have several explanations. Kezdi and Willis (2009) and Hudomiet et al. (2011) show that individual characteristics are strong determinants and that single households, in particular women, African-americans and less educated agents are more likely to be pessimistic about stock market returns and thus to hold fewer stocks. Hudomiet et al. (2011) also show that the market environment (and in particular the great recession of 2008) influences agents' expectations about stock market returns. Malmendier and Nagel (2011) look at the effect of having lived through the Great Depression. Agents who have experienced low stock market returns are also less likely to participate in the stock market and when they do, they are more pessimistic about future returns.

Financial education may nonetheless mitigate misperception (either pessimism or optimism) about the return of financial assets. Financial education is defined as "the process by which financial consumers/investors improve their understanding of financial products, concepts and risks and, through information, instruction and/or objective advice, develop the skills and confidence

[1] For a review of the possible explanations to the stock market participation puzzle, see Guiso and Sodini (2013).

to become more aware of financial risks and opportunities, to make informed choices, to know where to go for help, and to take other effective actions to improve their financial well-being" (OECD, 2005). Financial literacy or knowledge is then the outcome of this process. Lusardi et al. (2010) show that college education as well as parental (mother's) education and the financial sophistication of parents are important determinants. To the opposite, African-Americans, Hispanic, women, the young and the less-educated more often fail to understand financial literacy concepts (Lusardi and Mitchell, 2011). For this reason, public policies should primarily foster financial education of those more disadvantaged groups. Yet, as Kezdi and Willis (2009) argue, it takes "effort, intelligence and motivation to acquire knowledge of this body of evidence and use it to make saving and portfolio decisions that will raise the individual's or household's level of expected utility". As such, financial education is costly not only to governments but also to individuals.

It is clear that both misperception biaises and the lack of financial literacy have important lifetime economic consequences as they influence the size of accumulated wealth, the choice between different financial assets and ultimately, the possibility to smooth consumption across the lifecycle, in particular from the working-life period to the retirement period.[2] As mentioned in the OECD (2005) report, the acquisition of financial education and of financial knowledge has become crucial in recent years.[3] On the one hand, financial products have become more complex and more numerous. On the other hand, the changes in pension arrangements (i.e. the shift from defined benefits to defined contribution pension plans) as well as demographic changes (i.e. increased longevity and increased retirement period inducing lower replacement ratios) have shifted the responsibility of making lifetime financial planning from governments to individuals.[4]

Of course, if agents were perfectly farsighted, perfectly rational and perfectly informed about asset returns, no public intervention would be required and agents would be able to undertake the right decisions, i.e. the ones maximizing their expected utility. Since this is obviously not the case,

[2]Van Rooj et al. (2012) show a positive relationship between financial literacy and wealth accumulation. Financial knowledge reduces transaction costs (i.e. gathering and processing information) and is positively associated with retirement planning so that agents are better able to calculate their needs after retirement.

[3]We refer to financial education as the (costly) process through which individuals acquire financial literacy.

[4]On this, see also Lusardi and Mitchell (2007, 2014) and Lusardi et al. (2014).

the government may intervene through at least two mechanisms: public pensions and / or financial education. Historically, the first mechanism has been preferred and public pensions were meant to insure agents against the risk of having no (or too little) resources at retirement. However, because of increased financial pressure on public pension systems in most OECD countries, we have observed a shift from defined benefits (DB) retirement plans toward defined contributions (DC) plans, and therefore, a risk shifting from financial institutions to agents (e.g. future retirees).

When operating this shift from DB to DC plans, governments generally undertake additional measures to encourage the acquisition of financial literacy and to ensure that agents make appropriate financial decisions. One of these measures is providing or making mandatory financial education. For instance, in the Canadian context, Schwartz (2010) emphasizes that financial education should be a complement to changes in the pension structure. Lusardi and Mitchell (2011) also show that financial knowledge has a positive and statistically significant impact on retirement planning. In the same way, using data from Quebec, Lalime and Michaud (2014) show that there is a positive link between financial literacy, retirement preparation and the level of savings.

This paper precisely deals with the issue of the optimal mix between financial education programs and pension plans when agents have misperception biases regarding the expected return of their saving plans.[5] Misperception biases can take the form of either optimism or pessimism regarding the return of financial assets. The questions we ask are the following. Given the existence of misperception biases in the society, can the introduction of a DC plan make agents better-off? Equivalently, is the pension system a perfect remedy for the consequences of risks misperception on financial planning? If not, would financial education help solve these inefficiencies? To answer these questions, we model misperception biases as the combination of both an exogenous individual characteristic which accounts for optimism or pessimism and of financial education which can mitigate the impact of the initial misperception bias. We assume that agents are naive so that they never realize that their perception of the expected return on their assets is not correct. The government is utilitarian and paternalistic and thus maximizes the sum of the

[5] In this respect, our model is relevant for agents who are able to save. It does not apply to agents whose income is so low that they live hand-to-mouth.

true utilities, i.e. as if there were no misperception biases.[6] We solve the model assuming first that the government can perfectly observe individual biases. Second, we relax this assumption and assume more realistically that the government only knows the distribution of misperception biases in the society.

We obtain the following results. In the laissez faire, pessimistic agents invest too much in the riskless asset but too little in the risky one in comparison to the first-best optimum. The reverse is true for optimistic agents. As a result, the decentralization of the first-best optimum requires that pessimistic (resp. optimistic) agents face an individualized tax (resp. a subsidy) on the riskless asset and an individualized subsidy (resp. a tax) on the risky asset, together with individualized lump-sum transfers. We further assume that such a tax-and-transfer scheme is not available to the government, because, for example, it is impossible to observe savings or their allocation between risky and riskless assets. Instead, we assume the existence of a public pension system, such that agents must make mandatory contributions in the riskless asset for their retirement. We show that the level of contributions should be set equal to the amount of savings in the riskless asset that agents would choose if they were unbiased. Yet, such a system cannot alone solve all misperception problems. Optimistic agents get closer to the first-best solution, but they still invest too much in the risky asset. Furthermore, the introduction of the pension system is never a solution for pessimistic agents since it does not constrain their choices. They still end up at the laissez-faire solution. Hence, a pension scheme only constitutes a limited improvement of social welfare with respect to the laissez faire. We then allow for personalized levels of education in addition to the pension plan, and show that, because education is costly, it should be null for agents who are not too biased. However, beyond some threshold, education should increase with the size of the bias. We finally consider the more realistic case where the government does not observe individual biases and show how the optimal allocation is modified. We assume that agents are naive: they never realize that education has an impact on their misperception bias and thus, on their saving decisions, although it effectively has one. This proves to affect the design of the incentive compatibility constraints and of the second-best contracts in a non

[6] For a theorization of paternalistic social welfare functions, see Bernheim and Rangel (2007, 2009).

standard way. As a consequence, the single crossing condition does not hold. We show that, under asymmetric information, it is still optimal to set the pension benefit equal to the first best level of riskless investment. Education should always be distorted upward for every agent with a large misperception bias (whether optimistic or pessimistic). Agents with a low misperception bias should face a downward or an upward distortion of their education level. The direction of the distortion depends on the costs of education and on the distribution of the misperception bias in the economy.

The paper is organized as follows. The next section reviews the related literature. In Section 3, we describe the model, we derive the first best allocation and compare it with the laissez-faire. In Section 4, we focus on the case where the government can observe individual biases and we study to which extent a pension plan can help correcting misperception biases. Then, we introduce publicly provided education in this setting. In Section 5, we consider the case where the government cannot observe individual biases and has to rely on agents' self-reporting. The last section concludes.

2 Related literature

This paper builds on different strands of the literature. First, it is related to the prolific economic literature studying lifecycle decisions on consumption and savings.[7] These papers have shown how these decisions are affected by consumer preferences, economic shocks, public policies, etc. Yet, they generally assume that all agents are perfectly able to anticipate financial returns, ruling out any role for financial education.

Second, our paper contributes to the literature on the optimal design of pension systems and pension reforms. For instance, Matsen and Thorgensen (2004) have studied the optimal size of Pay-as-you-go pension systems allowing the agents to also invest in risky financial markets.[8] Yet, they assume perfect financial knowledge of agents/future retirees. Diamond and Koszegi (2003) and Cremer al. (2007, 2008, 2009) have also contributed to this literature by assuming that agents

[7]See among others, Ramsey (1928), Fisher (1930), Friedman (1957) and Samuelson (1958). For a review of the literature, see Browning and Crossley (2001).

[8]See also Dutta et al. (2000).

exhibit hyperbolic discounting. Conversely, in our model, agents correctly discount the future but misperceive financial returns, which does not only affects the overall amount of savings but also its composition. None of the existing papers in this literature considers financial education as a solution to such behavioral problems.

Third, we complement the growing theoretical literature studying the determinants and the consequences of financial education and financial literacy. As mentioned in Lusardi and Mitchell (2014), while the empirical literature is prolific (see the papers mentioned in the introduction), there are still few theoretical papers studying the role of financial education in correctly assessing financial returns. Some exceptions are Japelli and Padula (2013, 2015) who study, using a lifecycle model, the role of financial literacy and of financial education in explaining individual decisions (in terms of savings and literacy acquisition). They also show that pension benefits are substitutes to financial education as they reduce incentives to save and to accumulate financial literacy. Japelli and Padula (2015) consider the existence of two assets, a safe and a risky one. They find a positive correlation between financial education and stock holding and a negative correlation between social security generosity and stock holding.[9] More recently, Lusardi et al. (2017) simulate a multi-period dynamic lifecycle model to show that, in the United States, financial literacy can explain 30 to 40% of retirement wealth inequalities. As such, our paper is different from this strand of the literature, in that we adopt a normative approach and ask how much financial education agents *should* invest in, given that they do not perfectly evaluate financial returns. We go one step beyond by deriving the optimal combination between first-pillar pension benefits and financial education. Our study contributes to the policy debate at a time where worldwide pension reforms have recently transferred more responsibility regarding life-cycle saving decisions to individuals. Furthermore, except for Lusardi et al. (2017), the above models consider an homogeneous population while we assume a continuous distribution of misperception biases in the population and we derive optimal policies, both under symmetric and asymmetric information.

[9]See also Spataro and Corsini (2017) who provide a unified framework for studying the impact of financial education on human capital acquisition and on capital market participation (i.e. through risky and riskless assets). In their case, the decision to acquire financial knowledge is a 0-1 decision and is only relevant for investing in the risky asset. This is then more limited than our framework.

Finally, the most closely related paper is Corsini and Spataro (2014), who study the optimal individual choices between a safe and a risky pension plan as well as the optimal individual contribution rate to that pension plan (equivalent to a second pillar pension scheme), as a function of wage. They also assume that financial literacy decreases the cost of information related to choosing the risky plan. Their model differs from ours in several dimensions. First, contrary to us, they assume no mandatory first pillar of the pension system and assume instead that contributions to the complementary pension (either safe or risky) plan are a share the agent's wage. Second, we consider heterogeneity in the misperception biases, and imperfect information on the side of the government. Third, education in our model allows to better assess the true return of the risky asset and, as such, does not decrease the information costs related to a risky plan, which we assume to be null.

3 The model

This section first presents our assumptions. We then derive the first-best and the laissez-faire allocations, and show how the first-best optimum can be implemented.

3.1 Assumptions

Agents live two periods with certainty. In the first period, they supply a fixed amount of labor and earn an income w that they allocate between consumption and savings. The level of w is assumed to be identical across agents.[10] In the second period, they are retired and consume their savings. Agents can invest their savings in two assets: a riskless asset (a bond) with certain gross return r_b, and a risky asset (a stock) with gross return equal to r_s with probability π, and equal to 0 with probability $(1-\pi)$.[11] We denote by b the amount invested in the riskless asset, and by s the amount invested in the risky asset. In Sections 4 and 5, we allow for the existence of a public pension system, so that at, the time of retirement, agents obtain, in addition to their

[10]We assume no income heterogeneity to concentrate on the effects of misperception biases on the agents' investment choices and on how the government should intervene to correct for this inefficiency, putting aside any income redistribution concern.

[11]We could assume that the return of the risky asset in the bad state of nature is strictly positive. This would not qualitatively change our results.

private savings, a public pension benefit.

The individual utility function is linear in first-period consumption. In the second period, agents have a utility of consumption denoted by $u(.)$, with $u'(.) > 0$ and $u''(.) \leq 0$, and $u'(0) \to \infty$ (the Inada condition). The individual *true* (T) utility function is then written as follows:

$$U^T(c, d_1, d_2) = c + \beta[\pi u(d_1) + (1-\pi)u(d_2)],$$

where $c = w - b - s$ is first-period consumption, and $\beta \leq 1$ is the intertemporal discount factor. Second-period consumptions under the good state of nature (i.e. when the return on the risky asset is positive) and the bad state of nature (i.e. when the return on the risky asset is null) are respectively:

$$d_1 = sr_s + br_b, \quad (1)$$

$$d_2 = br_b. \quad (2)$$

Our model also assumes that agents misperceive the probability that the risky asset yields a positive return. They perceive this probability to be equal to $\phi(\alpha, e)\pi$ instead of π. The bias function $\phi(\alpha, e)$ (which we define formally below) depends on the agent's degree of myopia, α, which is different across agents, and on financial education, $e \in [0, 1]$. We assume a continuous distribution of types $\alpha \in [\alpha_{min}, 1/\pi]$ with $\alpha_{min} \in [0, 1]$. The density function is denoted by $f(\alpha)$ and the cumulative distribution function by $F(\alpha)$. We make the following assumption:

Assumption 1 $\alpha_{min} > r_b/r_s\pi$.

This ensures that, in the laissez faire, when no education is available (see Section 3.3), all agents invest in the risky asset (even though they may be very pessimistic).

The misperception bias has the following form

$$\phi(\alpha, e) = \alpha + (1-\alpha)e, \quad (3)$$

and is defined over $[0, 1/\pi]$ under our assumptions on the support of α and e. If $\alpha = 1$, the agent has the correct perception of the risk, and education would not change this. If $\alpha \neq 1$, the bias can be corrected by financial education e. If there is no financial education ($e = 0$), then the agent's bias is equal to α. If $\alpha < 1$, the agent is pessimistic about the expected return of the risky asset but education reduces the size of the bias so that $\alpha \leq \phi(\alpha, e) < 1$. If $\alpha > 1$, the agent is optimistic. Providing education reduces the size of the bias, so that $1 < \phi(\alpha, e) \leq \alpha$. Note also that $\phi_e(\alpha, e) = 1 - \alpha$ is positive (resp. negative) if α is smaller (resp. greater) than one so that education enables to converge toward the correct (unbiased) evaluation of the risk for both optimistic and pessimistic agents in a symmetric way. Finally, if education is perfect (i.e. $e = 1$), agents have no longer any bias.

Taking into account individual misperceptions, the individual *perceived* (P) utility function is written as:

$$U^P(c, d_1, d_2) = c + \beta \left[\phi(\alpha, e)\pi u(d_1) + (1 - \phi(\alpha, e)\pi)u(d_2)\right].$$

We make the assumption that agents are naive, thus ruling out self-control problems or sophistication. Agents believe they have the right perception of the risk, and that the good state of nature will realize with probability $\phi(\alpha, e)\pi$. Then, they do not understand the impact of education e on their perceptions and see no point in investing in it. For this reason, education can only be the result of public investment. We will introduce education in Sections 4 and 5.

We also assume a paternalistic planner, which implies that the social welfare function depends on the agents' *true* preferences and not their *perceived* ones. Such an assumption is particularly compelling since naive agents never realize they are biased.

3.2 First best optimum

The government is utilitarian and maximises the sum of a concave transformation of individual utilities. This amounts to solving the following problem:

$$\begin{aligned} \max_{c,b,s} W^{FB} &= \int_{\alpha_{min}}^{1/\pi} \Phi\{c + \beta\left[\pi u(d_1) + (1 - \pi)u(d_2)\right]\} f(\alpha) d\alpha \\ \text{s.t.} \quad & \int_{\alpha_{min}}^{1/\pi} (w - c - b - s) f(\alpha) d\alpha \geq 0, \end{aligned}$$

where $\Phi(.)$ is an increasing and concave transformation of the individual utility, and d_1 and d_2 are defined by equations (1) and (2) respectively. The second line in the above problem is the resource constraint of the economy. The first order conditions (FOCs hereafter) are:

$$\frac{\partial W}{\partial c} = \Phi'(U^T) - \lambda = 0, \tag{4}$$

$$\frac{\partial W}{\partial b} = \Phi'(U^T)\beta r_b \left[\pi u'(d_1) + (1-\pi)u'(d_2)\right] - \lambda = 0, \tag{5}$$

$$\frac{\partial W}{\partial s} = \Phi'(U^T)\beta r_s \pi u'(d_1) - \lambda = 0, \tag{6}$$

where λ is the Lagrange multiplier associated with the resource constraint. These conditions together give the first best levels of the risky and riskless investments, which we denote s^{FB} and b^{FB}. It is straightforward to see that $U^T(c, d_1, d_2) = \bar{U}\ \forall\alpha$. Rearranging the FOCs, we obtain that $d_1^{FB} = \bar{d}_1$ and $d_2^{FB} = \bar{d}_2$, for all α. This also implies that first-period consumption is the same for every agent irrespective of their perception bias, i.e. $c = \bar{c}\ \forall\alpha$. Furthermore, we obtain the following relationship between consumption levels:

$$\frac{u'(\bar{d}_2)}{u'(\bar{d}_1)} = \frac{\pi}{1-\pi}\frac{r_s - r_b}{r_b}\ \forall\alpha. \tag{7}$$

Under Assumption 1, the RHS of this equation is always greater than 1, implying that, in the first best, second-period consumption in the good state of nature, $\bar{d}_1$, is greater than consumption in the bad state of nature, $\bar{d}_2$, for every agent.

3.3 Laissez faire allocation

Let us then study the laissez-faire allocation and compare it with the first-best optimum.

In the laissez faire, there is no government intervention so that $e = 0$ and the individual misperception bias simplifies to $\phi(\alpha, 0) = \alpha$. Each agent with bias α maximizes the following (perceived) utility function

$$U^P(s,b) = w - s - b + \beta\left[\alpha\pi u(sr_s + br_b) + (1-\alpha\pi)u(br_b)\right]$$

with respect to b and s, under the constraints $b \geq 0$ and $s \geq 0$. The first-order conditions are, respectively:

$$\frac{\partial U^P}{\partial b} = -1 + \beta r_b \left[\alpha\pi u'(d_1) + (1-\alpha\pi)u'(d_2)\right] \leq 0, \quad (8)$$

$$\frac{\partial U^P}{\partial s} = -1 + \beta\alpha\pi r_s u'(d_1) \leq 0. \quad (9)$$

Under the Inada condition, b is always strictly positive in order to ensure positive consumption in the bad state of the world (i.e. if the risky asset yields a zero return). Thus, the first condition always holds with equality and substituting this condition into equation (9), we get

$$\frac{\partial U^P}{\partial s} = \beta\left[-(1-\alpha\pi)r_b u'(d_2) + (r_s - r_b)\alpha\pi u'(d_1)\right] \leq 0.$$

Evaluating it in $s = 0$, we find that $s > 0$ if and only if $r_s\pi\alpha > r_b$. Under Assumption 1, this is always the case, and every agent chooses to invest in the risky asset. Then, the first-order condition with respect to s always holds with equality. Combined with (8), (9) yields:

$$\frac{u'(d_2)}{u'(d_1)} = \frac{\alpha\pi}{(1-\alpha\pi)}\frac{r_s - r_b}{r_b} > 1 \ \forall\alpha. \quad (10)$$

For further use, let us define $b^{LF}(\alpha)$ and $s^{LF}(\alpha)$, the laissez faire saving choices of an agent with bias α. They are obtained by solving equations (8) and (9). We show in the appendix that:

$$\frac{\partial b^{LF}(\alpha)}{\partial\alpha} \leq 0, \quad (11)$$

$$\frac{\partial s^{LF}(\alpha)}{\partial\alpha} \geq 0. \quad (12)$$

This also implies that d_1 increases with α while d_2 decreases with it.

Let us then compare the laissez-faire allocation with the first best optimum. The RHS of (10) and (7) differ by the parameter α. If agents have the correct perception of the risk, (i.e. $\alpha = 1$) they take unbiased decisions, and the laissez-faire allocation is first-best optimal. Using (11) and (12), it is clear that the laissez faire is in general not first-best optimal, as we summarize it in the following proposition:

Proposition 1 *Consider an economy where agents invest both in a risky and a riskless asset, and differ in their levels of misperception α of the risk associated to the risky asset. At the laissez faire,* *i) Optimistic agents (i.e. with $\alpha > 1$) overinvest in the risky asset and underinvest in the riskless asset with respect to the first best optimum: $s^{LF}(\alpha) > s^{FB} = s^{LF}(1)$ and $b^{LF}(\alpha) < b^{FB} = b^{LF}(1)$*

ii) Pessimistic agents (i.e. with $\alpha < 1$) underinvest in the risky asset and overinvest in the riskless asset with respect to the first best optimum: $s^{LF}(\alpha) < s^{FB} = s^{LF}(1)$ and $b^{LF}(\alpha) > b^{FB} = b^{LF}(1)$

iii) Non biased agents (i.e. with $\alpha = 1$) choose the optimal levels of risky and riskless assets.

As we have just shown, the intervention of the government is then required to correct for inefficient individual decisions. In order to decentralize the first best, the government needs to set an individualized tax on the risky asset, $\tau(\alpha)$, as well as an individualized tax, $\theta(\alpha)$, on the riskless asset. To ensure a uniform level of first-period consumption (and thus, equal expected utilities), the government also needs to use personalized lump sum taxes, $T(\alpha)$. We show in Appendix 6, that the optimal taxes are

$$\begin{aligned} \tau^*(\alpha) &= 1 - \frac{1}{\alpha}, \\ \theta^*(\alpha) &= 1 - [\frac{\pi u'(\bar{d}_1) + (1-\pi)u'(\bar{d}_2)}{\alpha\pi u'(\bar{d}_1) + (1-\alpha\pi)u'(\bar{d}_2)}]. \end{aligned}$$

Agents with $\alpha > 1$ are taxed when investing in the risky asset while those with $\alpha < 1$ are subsidized. Using the first-best result that $u'(\bar{d}_1) < u'(\bar{d}_2)$, we also find that every agent with $\alpha < 1$ faces a tax on the riskless asset (i.e. $\theta(\alpha) > 0$), while agents with misperception $\alpha > 1$ face a subsidy on the non risky asset (i.e. $\theta(\alpha) < 0$). These results are intuitive. In the laissez faire, agents with $\alpha < 1$ buy too much of the riskless asset and not enough of the risky one. It is then optimal to tax the riskless asset and to subsidize the risky one. The reverse is true for pessimistic agents. Agents with $\alpha = 1$ have the correct perception of financial and are not taxed.

The optimal levels of lump sum taxes are equal to:

$$T(\alpha) = w - \bar{c} - \frac{b^{LF}(1)}{1-\theta^*(\alpha)} - \frac{s^{LF}(1)}{1-\tau^*(\alpha)} \gtrless 0, \tag{13}$$

where $\bar{c}$ is the first-best level of first-period consumption (set such that the resource constraint of the economy is satisfied).

4 Pension system under symmetric information

In reality, it may be difficult for a government to implement the linear taxes on (either risky or riskless) investments we studied in the previous section. This may be the case in particular when total saving amounts and/or the decomposition between risky and riskless assets are difficult to observe. This is why, in this section, we study alternative settings in which the government cannot implement such a tax-and-transfer scheme. Instead, we are going to assume first that the government can implement a mandatory pension system (Section 4.1) to ensure higher resources in the retirement period and partially cope for inefficient planning due to misperception biases.[12] Second, we assume that, in addition to the pension plan, the government can introduce personalized financial education (Section 4.2). Our objective is to study how these policy instruments can increase aggregate welfare when agents exhibit misperception biases, assuming that the government perfectly observes these biases. We will consider asymmetric information in Section 5.

4.1 Pension system with no education

In the first period, agents pay mandatory pension contributions for an amount $B(\alpha)$ invested in the riskless asset. In the second period, they get retirement benefits equal to $r_b B(\alpha)$. Under the reasonable assumption that the pension system has to be balanced *ex post*, individual contributions to the pension system are never invested in the risky asset.[13] We also assume that, in order to redistribute across agents, the government implements individualized lump sum taxes $T(\alpha)$.

The timing is the following. The government announces the fiscal policy, that is $\{B(\alpha), T(\alpha)\}$, in the first stage. In the second stage, agents choose their investment levels in the riskless and risky assets, taking as given the policy instruments. We assume that agents cannot borrow against future pensions so that $b(\alpha) \geq 0$ and $s(\alpha) \geq 0$ for all α. As usual in this type of problem, we

[12] This is equivalent to the first pillar of the pension system.

[13] We will discuss the case where part of the pension contribution is invested in the risky asset in the concluding section.

proceed backward. First, we derive the agents' choices as a function of their bias and of the policy instruments. Second, we determine the optimal levels of $B(\alpha)$ and $T(\alpha)$ for every agent of type α.

Introducing $T(\alpha)$ and $B(\alpha)$, the individual problem is now[14]

$$\max_{s,b} U^P(s,b) = w - s - b - B - T + \beta\left[\alpha\pi u(sr_s + (b+B)r_b) + (1-\alpha\pi)u((b+B)r_b)\right].$$

The FOCs with respect to b and s of an agent with bias α are:

$$\frac{\partial U^P}{\partial b} = -1 + r_b\beta[\alpha\pi u'(d_1) + (1-\alpha\pi)u'(d_2)] \leq 0, \tag{14}$$

$$\frac{\partial U^P}{\partial s} = -1 + r_s\beta\alpha\pi u'(d_1) \leq 0 \tag{15}$$

where d_1 and d_2 now include the pension contribution $B(\alpha)$. These conditions give the individual choices in terms of riskless and risky assets, denoted by $b(\alpha, B)$ and $s(\alpha, B)$ respectively. Because preferences are quasi-linear, individual choices are independent of the level of the lump sum tax $T(\alpha)$. Regarding the levels of $b(\alpha, B)$ and $s(\alpha, B)$, two cases can arise depending on the comparison between the levels of $B(\alpha)$ and $b^{LF}(\alpha)$.

If $B(\alpha) \leq b^{LF}(\alpha)$, the agent chooses to top up the pension contribution $B(\alpha)$ by further investing in the riskless asset. Hence, the investment in the riskless asset when there exists a pension system is $b(\alpha, B) = b^{LF}(\alpha) - B(\alpha)$, and $\partial b(\alpha, B)/\partial B = -1$. Since total investment in the riskless asset remains the same as in the laissez faire, the investment in the risky asset is not affected by the existence of the pension plan, so that $s(\alpha, B) = s^{LF}(\alpha)$ and $\partial s(\alpha, B)/\partial B = 0$.

If $B(\alpha) > b^{LF}(\alpha)$, agents would actually prefer a lower level of investment in the riskless asset than the level obtained through the mandatory pension plan. As a result, they set $b(\alpha, B) = 0$, so that $\partial b(\alpha, B)/\partial B = 0$. The FOC with respect to $s(\alpha, B)$, (15), simplifies to

$$-1 + \pi\alpha r_s u'(sr_s + Br_b) = 0.$$

Fully differentiating this condition, we also find that

$$\frac{\partial s(\alpha, B)}{\partial B} = -\frac{r_b}{r_s} < 0,$$

[14]For ease of notation, we drop the arguments of the functions whenever this does not lead to ambiguity.

so that as $B(\alpha)$ increases, $s(\alpha, B)$ decreases. This result is not surprising. Since agents are pushed to invest more in the riskless asset than what they would want to, their marginal utility in the good state of nature, $u'(d_1)$, is lower and they prefer to reduce their investment in the risky asset.

Taking into account the individual responses, in the second stage the paternalistic government maximizes the social welfare function subject to the resource constraint:

$$\begin{aligned} \max_{B(\alpha),T(\alpha)} W = & \int_{\alpha_{min}}^{1/\pi} \Phi\{w - s - b - B(\alpha) - T(\alpha) \\ + & \beta \left[\pi u(sr_s + (b + B(\alpha))r_b) + (1-\pi)u((b + B(\alpha))r_b)\right]\} f(\alpha)d\alpha \\ \text{s.t.} & \int_{\alpha_{min}}^{1/\pi} T(\alpha)f(\alpha)d\alpha \geq 0, \end{aligned}$$

where b and s are obtained from (14) and (15). In what follows, we assume that the pension contribution cannot be negative, $B(\alpha) \geq 0 \; \forall \alpha$. The FOCs with respect to $B(\alpha)$ and $T(\alpha)$, for each α type are [15]

$$\begin{aligned} \frac{\partial W}{\partial B(\alpha)} &= f(\alpha)\Phi'(U^T) \times \left\{ \{-1 + \beta r_b \left[\pi u'(d_1) + (1-\pi)u'(d_2)\right]\} \left(1 + \frac{\partial b}{\partial B}\right) \right. \\ & \left. + \{-1 + \beta \pi r_s u'(d_1)\} \frac{\partial s}{\partial B} \right\} \leq 0, \quad (16) \\ \frac{\partial W}{\partial T(\alpha)} &= f(\alpha)(-\Phi'(U^T) + \lambda) = 0, \quad (17) \end{aligned}$$

where λ is the Lagrange multiplier associated with the government's budget constraint. As in the first best (see Section 3.2), condition (17) states that $T(\alpha)$ should be set such that true indirect expected utility levels are equalized across all agents.

Let us then study condition (16). To do so, we evaluate the RHS of (16) in $B(\alpha) = b^{LF}(\alpha)$. If $B(\alpha)$ further increases above that level, then $\partial b/\partial B = 0$ and $\partial s/\partial B \leq 0$. Then, substituting for (14) and (15) (set to equality), the first term in (16) is negative if $\alpha < 1$, and positive if $\alpha > 1$. Similarly, the second term in (16) is negative if $\alpha < 1$, and positive if $\alpha > 1$.

Hence, for optimistic agents ($\alpha > 1$), $\partial W/\partial B(\alpha)|_{B(\alpha)=b^{LF}(\alpha)} > 0$ so that the optimal contribution $B^O(\alpha)$ should be greater than the laissez-faire riskless investment level, $b^{LF}(\alpha)$. Indeed, in the laissez faire, these agents underinvest in b and overinvest in s, so that a mandatory contribution into the riskless asset improves their welfare. As a result, they will choose $b(\alpha, B^O(\alpha)) = 0$.

[15] We assume that the second-order conditions hold.

Moreover, differentiating (16) with respect to α and using the implicit function theorem, we obtain that $\partial B^O(\alpha)/\partial\alpha$ has the same sign as

$$\beta\pi r_s u''(d_1)\frac{\partial s}{\partial\alpha}\left(\frac{\partial s}{\partial B}r_s + r_b\right) + \{-1 + \beta\pi r_s u'(d_1)\}\frac{\partial^2 s}{\partial B\partial\alpha}.$$

Recalling that $\partial s/\partial B = -r_b/r_s$, the above equation is equal to zero so that $\partial B^O(\alpha)/\partial\alpha = 0$. This implies that the optimal level of contributions for optimistic agents does not depend on α. From eq. (20), the optimal level of the uniform contribution for optimistic agents satisfies

$$\begin{aligned}\frac{\partial W}{\partial B} &= f(\alpha)\Phi'(U^T) \times \{-1 + \beta r_b\left[\pi u'(d_1) + (1-\pi)u'(d_2)\right] - [-1 + \beta\pi r_s u'(d_1)]\frac{r_b}{r_s}\} = 0 \\ &= f(\alpha)\Phi'(U^T) \times \{-1 + \frac{r_b}{r_s} + \beta(1-\pi)r_b u'(d_2)\} = 0\ \ \forall\alpha.\end{aligned} \tag{18}$$

Replacing (4) and (6) into (5), one can easily see that the condition above and condition (5) are identical. We then obtain that $B^O(\alpha)$ should optimally be set at b^{FB} for all $\alpha > 1$.

Here, the government has only one instrument (i.e. the mandatory contributions in the riskless asset) to correct for two inefficiencies: underinvestment in the riskless asset and overinvestment in the risky asset. With mandatory contributions $B^O(\alpha) = b^{FB}$ invested in the riskless asset, optimistic agents are forced to invest the optimal amount in the riskless asset (they will now do so exclusively through the pension system). In addition, as shown by the second term in (16), by setting the level of the public contribution, the government will indirectly partially correct for the overinvestment in the risky asset. Optimistic agents will lower their investment s (recall that s and B are substitutes) and make it closer to its first-best level. Increasing B above its first-best level would further decrease the overinvestment in the risky asset, but this would imply overinvesting in the riskless asset as well. All in all, the government policy enhances the welfare of optimistic agents with respect to the laissez faire but does not fully correct for the consequences of the misperception bias.

Let us now consider pessimistic agents (i.e. $\alpha < 1$). Their optimal contribution $B^O(\alpha)$ should be smaller than or equal to $b^{LF}(\alpha)$, since, at the laissez faire, these agents overinvest in b. Thus, a retirement contribution invested in the riskless asset higher than their laissez-faire investment in that asset can never be optimal. In addition, any level of $B^O(\alpha)$ smaller than $b^{LF}(\alpha)$ yields the same welfare as in the laissez faire: agents will simply top-up the retirement contribution with their voluntary savings so as to exactly invest $b^{LF}(\alpha) = B^O(\alpha) + b(\alpha, B^O(\alpha))$ in the riskless asset, and choose accordingly $s^{LF}(\alpha)$. Thus, any contribution level, $B^O(\alpha) \in \left[0, b^{LF}(\alpha)\right]$ has no impact on the investment decisions of pessimistic agents as well as on their utility.[16] The pension system is therefore not a solution to the behavioral bias of pessimistic agents.

Our results are summarized in the following proposition:

Proposition 2 *Consider an economy where agents invest both in a risky and a riskless asset, and differ in their levels of misperception α of the risk associated to the risky asset. If a paternalistic government observes individual biases and can only intervene through a mandatory pension system,*

i) For optimistic agents (i.e. with $\alpha \geq 1$), the optimal contribution is $B^O(\alpha) = b^{FB} = b^{LF}(1)$. Agents do not invest privately in the riskless asset and adapt accordingly their investment in the risky asset, $s^{LF}(\alpha) > s^O(\alpha)$.

ii) For pessimistic agents (i.e. with $\alpha < 1$), the optimal contribution is $B^O(\alpha) \in [0, b^{LF}(\alpha)]$. They complement with investment in the riskless asset so as to obtain exactly $b^{LF}(\alpha)$. Investment in the risky asset is the same as in the laissez faire.

Given these results, it is then optimal for the government to set a *uniform* level of contributions, $b^{FB} = b^{LF}(1)$ for all agents with different α. Interestingly, such a solution does not require any information on the agents' bias.

Nonetheless, we find that introducing a pension system alone cannot solve all misperception problems. Optimistic agents get closer to the first best solution, but the introduction of the pension system is not enough to make them invest the right amount in the risky asset. It is also never a solution for pessimistic agents since it does not constrain their choices and they end up at

[16] If public funds were costly, the optimal contribution would be equal to zero, as the same (laissez-faire) welfare level could be attained either through B or b but at a smaller cost for the society in the latter case.

the laissez-faire solution. In the next section, we study how publicly provided financial education, in addition to a pension plan, could help solving these inefficiencies.

4.2 Pension system with financial education

In this section, we introduce an additional policy instrument: a personalized mandatory level of financial education in the first period, $e(\alpha)$, set by the government. Here, we assume that the government can observe the individual behavioral biases and we look at whether education, together with the pension scheme, increases social welfare.

The timing is the same as in the previous section so that, as above, we first solve the individual problem, for a given menu of policy instruments $\{B(\alpha), e(\alpha), T(\alpha)\}$, and then derive the government's problem.

We assume that education entails a linear disutility cost, $pe(\alpha)$ to the agent, where p is the unit private cost of education. The utility of an agent with bias α is modified as follows

$$U^P(s,b) = w - s - b - pe - B - T + \beta\left[\phi(\alpha,e)\pi u(sr_s + (b+B)r_b) + (1-\phi(\alpha,e)\pi)u((b+B)r_b)\right],$$

where $\phi(\alpha, e)$ is defined by (3). The FOCs with respect to b and s are, respectively:

$$\frac{\partial U^P}{\partial b} = -1 + r_b\beta[\phi(\alpha,e)\pi u'(d_1) + (1-\phi(\alpha,e)\pi)u'(d_2)] \le 0, \tag{19}$$

$$\frac{\partial U^P}{\partial s} = -1 + r_s\beta\phi(\alpha,e)\pi u'(d_1) \le 0, \tag{20}$$

where, as before, d_1 and d_2 include the pension benefit. These equations define the investments in the risky and riskless assets as a function of the policy instruments and of the agent's type: $s(\alpha, B, e)$ and $b(\alpha, B, e)$.[17]

When both $s(\alpha, B, e) > 0$ and $b(\alpha, B, e) > 0$, we show in the Appendix that

$$\frac{\partial b(\alpha,B,e)}{\partial e} \begin{cases} < 0 & if\ \alpha < 1 \\ \ge 0 & if\ \alpha \ge 1 \end{cases}$$

and

$$\frac{\partial s(\alpha,B,e)}{\partial e} \begin{cases} > 0 & if\ \alpha < 1 \\ \le 0 & if\ \alpha \ge 1. \end{cases}$$

[17] As before, due to the quasi-linearity of preferences, individual decisions do not depend on the level of $T(\alpha)$.

Intuitively, for a given level of $B(\alpha)$, increasing the investment in education e leads to an increase in s and a decrease in b for pessimistic agents and to a decrease in s and an increase in b for optimistic agents. We also show in the Appendix that, as in the previous section, whenever α increases, $b(\alpha, B, e)$ decreases and $s(\alpha, B, e)$ increases.

Let us now derive the problem of the paternalistic government. It consists in solving

$$\begin{aligned}\max_{B(\alpha),e(\alpha),T(\alpha)} W = \quad & \int_{\alpha_{min}}^{1/\pi} \Phi\{w - s - b - B(\alpha) - pe(\alpha) - T(\alpha) \\ & +\beta\left[\pi u(sr_s + (B(\alpha)+b)r_b) + (1-\pi)u((B(\alpha)+b)r_b)\right]\}f(\alpha)d\alpha\end{aligned}$$

subject to the resource constraint of the government

$$\int_{\alpha_{min}}^{1/\pi} T(\alpha)f(\alpha)d\alpha \geq q\int_{\alpha_{min}}^{1/\pi} e(\alpha)f(\alpha)d\alpha,$$

where q is the unit public cost of education and where s and b are obtained from FOCs (19) and (20).

The FOCs of the government problem are

$$\begin{aligned}\frac{\partial W}{\partial B(\alpha)} &= \{-1+\beta r_b\left[\pi u'(d_1)+(1-\pi)u'(d_2)\right]\}\left(1+\frac{\partial b}{\partial B}\right) \\ &+ \{-1+\beta\pi r_s u'(d_1)\}\frac{\partial s}{\partial B} \leq 0, \qquad (21)\\ \frac{\partial W}{\partial e(\alpha)} &= -(p+q)+\{-1+\beta r_b\left[\pi u'(d_1)+(1-\pi)u'(d_2)\right]\}\frac{\partial b}{\partial e} \\ &+ \{-1+\beta\pi r_s u'(d_1)\}\frac{\partial s}{\partial e} \leq 0, \qquad (22)\\ \frac{\partial W}{\partial T(\alpha)} &= -\Phi'(U^T)+\lambda = 0, \qquad (23)\end{aligned}$$

where λ is the Lagrange multiplier associated with the government's budget constraint.

As in the previous section, $T(\alpha)$ is set so as to equalize the expected true utilities of agents with different α. The FOC with respect to $B(\alpha)$, (21), is identical to the one with no education, (16), so that it can be analyzed exactly in the same way. For optimistic agents, it is optimal to set the level of mandatory contributions to $b^{FB} = b^{LF}(1)$. For pessimistic agents, such a pension contribution has no effect, since they top up the contributions with extra investment in the riskless asset, and therefore end up with the laissez-faire allocation. In the following and without loss of generality, we therefore set $B^*(\alpha) = b^{FB} = b^{LF}(1)$ for every (optimistic or pessimistic) agent, where a * denotes the optimal level of B when education is also publicly provided.

Let us now study the condition on financial education. Using the agent's FOCs (19) and (20), it can be rewritten as

$$\begin{aligned}\frac{\partial W}{\partial e} = -(p+q) + \{\beta r_b(1-\phi(\alpha,e))\pi\left[u'(d_1) - u'(d_2)\right]\}\frac{\partial b}{\partial e} \\ + \{\beta\pi r_s(1-\phi(\alpha,e))u'(d_1)\}\frac{\partial s}{\partial e} \leq 0 \ \ \forall\alpha. \qquad (24)\end{aligned}$$

If the paternalistic government sets $e = 1$, $\phi(\alpha, e) = 1$ and there is no misperception any longer. Yet, investing in education entails costs (both public and private) in terms of first-period consumption, so that it is optimal to have $e(\alpha) < 1$, $\forall\alpha$.[18] For that reason, it is optimal to implement both a pension system, which is at no cost for the economy, and financial education.

By fixing the contribution to $B^*(\alpha) = b^{FB}$, optimistic agents are constrained to invest in the riskless asset more than they wish (even if education lowers their misperception bias) and so they choose $b(\alpha, b^{FB}, e) = 0$. For them, education only has an impact on the level of the risky asset, $s(\alpha, b^{FB}, e)$. Conversely, for pessimistic agents, who end up with their laissez-faire allocation when $B^*(\alpha) = b^{FB}$, financial education is useful to correct both overinvestment in the riskless asset and underinvestment in the risky asset.

Using the comparative statics on $b(\alpha, b^{FB}, e)$ and $s(\alpha, b^{FB}, e)$ derived above, we obtain that the second term in (24) is positive if $\alpha < 1$ and null if $\alpha > 1$, while the third term is always positive. These terms represent the benefits of education on $b(\alpha, b^{FB}, e)$ and $s(\alpha, b^{FB}, e)$, respectively. The first term in (24) represents the social marginal cost of education and it is always negative.

When $\alpha = 1$, equation (24) reduces to $-(p + q)$ and the optimal level of financial education is zero. This is a direct consequence of the agent not being biased. In this case, education only creates a cost but no benefit. Since, for all $p + q > 0$, equation (24) evaluated in $\alpha = 1$ and $e = 0$ is strictly negative, there exist two thresholds $\{\alpha_1; \alpha_2\}$ with $\alpha_1 < 1 < \alpha_2$ such that the optimal level of education is constrained to zero for agents with $\alpha \in [\alpha_1, \alpha_2]$.[19] The intuition here is that,

[18] It is straightforward to see that (24) evaluated at $e = 1$ is negative.

[19] The thresholds are obtained such that (24) evaluated in $\{e = 0, \alpha = \alpha_1\}$ or in $\{e = 0, \alpha = \alpha_2\}$ is equal to zero.

for low levels of misperception (no matter whether pessimistic or optimistic), the marginal cost of correcting the bias through education is higher than its marginal benefit. Hence, agents need to be "sufficiently" biased for the government to intervene.

We obtain the following results regarding the optimal government intervention:

Proposition 3 *Consider an economy where agents invest both in a risky and a riskless asset, and differ in their levels of misperception α of the risk associated to the risky asset. If a paternalistic government observes individual biases and can intervene through a pension system and personalized levels of financial education, the optimal policy consists in*

i) a uniform pension contribution $B^*(\alpha) = b^{FB} = b^{LF}(1)\ \forall\alpha$,

ii) a zero level of financial education $e^*(\alpha) = 0$ for agents with $\alpha \in [\alpha_1, \alpha_2]$,

iii) a positive level of financial education otherwise. If preferences display constant absolute risk aversion (CARA), $e^*(\alpha)$ is decreasing in α whenever $\alpha < \alpha_1$, and increases in α whenever $\alpha > \alpha_2$.

Proof. For points *i)* and *ii)*, see above. For point *iii)*, see the Appendix.[20] ■

To understand the above proposition, consider two agents who are both pessimistic. According to our results, the more pessimistic one should receive more education than the less pessimistic one (who is closer to the right perception). Similarly, consider two optimistic agents. The more optimistic one should receive more education than the less optimistic one. Note however that an optimistic and a pessimistic agents whose bias is the same in absolute size will not receive the same amount of education. This is due to the fact that pessimistic agents misinvest in both assets, while optimistic agents invest the right amount in the riskless asset (b^{FB}).

We now turn to studying the optimal lump sum tax, $T(\alpha)$. Equation (23) shows that true indirect utilities (which only take into account the misperception bias through saving decisions

[20]When preferences are not CARA, we cannot obtain unambiguous results. Indeed, with CARA preferences, the bias α affects risk taking only directly through $\phi(\alpha, e)$, but not indirectly through its impact on marginal utilities of consumption, $u'(.)$.

and the policy instruments) should be equalized at the first best. This implies that

$$\frac{dU^T}{d\alpha} = \{-1+\beta r_b\left[\pi u'(d_1)+(1-\pi)u'(d_2)\right]\}\frac{db}{d\alpha} + \{-1+\beta\pi r_s u'(d_1)\}\frac{ds}{d\alpha} - p\frac{de}{d\alpha} - \frac{dT(\alpha)}{d\alpha} = 0,$$

which leads to

$$\frac{dT(\alpha)}{d\alpha} = \{\beta r_b(1-\phi(\alpha,e))\pi\left[u'(d_1)-u'(d_2)\right]\}\frac{db}{d\alpha} + \{\beta\pi r_s(1-\phi(\alpha,e))u'(d_1)\}\frac{ds}{d\alpha} - p\frac{de}{d\alpha}, \tag{25}$$

where we made use of the FOCs on $b(\alpha, b^{FB}, e)$ and $s(\alpha, b^{FB}, e)$. As before, note that $b(\alpha, b^{FB}, e) = 0$ for optimistic agents. Assuming that $db/d\alpha$ has the same sign as $\partial b/\partial\alpha$ (which has been shown to be negative in the Appendix) and that $ds/d\alpha$ has the same sign as $\partial s/\partial\alpha$ (which has been shown to be positive in the Appendix), and using the variation of e with α (see Proposition 3), we obtain the following result:[21]

$$\frac{dT(\alpha)}{d\alpha} \begin{cases} > 0 \text{ for } \alpha < 1 \\ \leq 0 \text{ for } \alpha \geq 1, \end{cases}$$

so that $T(\alpha)$ is inversed U-shaped in α. Let us also study how the individual total contribution $T(\alpha)+pe(\alpha)$ varies with the level of misperception α. Using (25), it is straightforward to see that

$$\frac{d(T(\alpha)+pe(\alpha))}{d\alpha} \begin{cases} > 0 \text{ for } \alpha < 1 \\ \leq 0 \text{ for } \alpha \geq 1. \end{cases}$$

Hence, at the extremes of the α distribution, the level of education is relatively high, but the tax is relatively low (possibly negative), while for agents with a low degree of misperception (close to α_1 and α_2), the level of education is low but taxation is important. Overall, $T(\alpha) + pe(\alpha)$ is increasing up to $\alpha = 1$, and then decreasing.

This implies that, under asymmetric information, the full information policy would not be incentive compatible: agents with low degrees of misperception would always mimic the most

[21] The assumption that $db/d\alpha$ and $\partial b/\partial\alpha$ have the same sign is equivalent to assuming that even with education, $\phi(\alpha, e(\alpha))$ increases in α. The same is true for the assumption that $ds/d\alpha$ and $\partial s/\partial\alpha$ have the same sign.

biased agents. In the following section, we assume asymmetric information and show how the government's problem is modified by the introduction of incentive compatibility constraints.

5 Asymmetric information

5.1 The individual problem and incentive compatible mechanisms

Let us now turn to the case where the government observes the distribution of α but not individual biases, and consequently has to propose an incentive compatible allocation of pension benefits, education and taxes, such that agents truthfully reveal their type.

Since agents are naive, they never realize how biased they are. They do not understand the impact of e on $\phi(\alpha, e)$, even tough it has one. Before benefiting from education, agents simply believe they will obtain a high return with probability $\rho = \phi(\alpha, 0)\pi = \alpha\pi$. Since, by definition, naive agents are not aware of their α and can only assess the perceived probability, ρ, this is the variable they report to the government under a direct revelation mechanism.[22] In the following, we thus make a change of variable and rewrite α as ρ/π, with $\rho \in [\pi\alpha_{min}, 1]$ and density function $f_\rho(\rho) = f(\rho/\pi)/\pi$.[23]

The government will then propose a menu of incentive compatible contracts, $\{B, e(\rho), T(\rho)\}\ \forall\rho$. As before, financial education, $e(\rho)$, is personalized but the pension contribution B will be uniform.[24] As a consequence, the mimicking behaviour only depends on the variation of $T(\rho)+pe(\rho)$, as we showed at the end of Section 4.2.

The timing is the same as before. In the first stage, the government proposes a contract $(B, e(\rho), T(\rho))\ \forall\rho$. In the second stage, agents report ρ. Under the assumption that agents are biased and naive, the level of education enters their utility only as a cost, since they do not realize it has an impact on their investment choices. In the third stage, agents choose the investment levels in the risky and riskless assets. As usual, we proceed by backward induction.

[22] Since the revelation theorem applies in this context, we focus on a direct mechanism.

[23] The cumulative density function $F_\rho(\hat{\rho}) = P(\rho \leq \hat{\rho}) = P(\alpha\pi \leq \hat{\rho}) = F(\rho/\pi)$. Differentiating this expression, we obtain that the density function of ρ is equal to $f(\rho/\pi)/\pi$.

[24] As we showed in the previous section, under full information it is indeed optimal to set the same (first-best) level of mandatory contributions. To keep the model treatable and the results comparable with the previous ones, we limit the analysis to uniform benefits also in this section, even if they may be suboptimal under asymmetric information.

Since agents are naive, we will need, in the following, to distinguish between what we will call *ex-post* and *ex-ante* expected utilities. The ex-ante utility will be useful in computing incentive compatibility constraints while the ex-post utility will be useful when deriving the government's problem.

Let us start with the *ex-post* (EP) expected utility. It is the individual expected utility after the bias has been partially corrected by financial education, e:

$$U_\rho^{EP}(s,b) = w - s - b - pe - T - B$$
$$+ \beta[\phi(\rho/\pi,e)\pi u(sr_s + (b+B)r_b) + (1-\phi(\rho/\pi,e)\pi)u((b+B)r_b)].$$

Agents are still biased but their decisions are influenced by education. The investment levels are given by $s^{EP} \equiv s(\phi(\rho/\pi,e),B,e)$ and $b^{EP} \equiv b(\phi(\rho/\pi,e),B,e)$, which maximize the utility above and are implicitly defined as follows:

$$-1 + r_b\beta[\phi(\rho/\pi,e)\pi u'(d_1^{EP}) + (1-\phi(\rho/\pi,e)\pi)u'(d_2^{EP})] \leq 0, \tag{26}$$

$$-1 + r_s\beta\phi(\rho/\pi,e)\pi u'(d_1^{EP}) \leq 0, \tag{27}$$

where $d_1^{EP} = r_b(b^{EP}+B) + r_s s^{EP}$, and $d_2^{EP} = r_b(b^{EP}+B)$.

Let us now analyze how the levels b^{EP} and s^{EP} vary with the level of mandatory contributions. To do so, we define $b(\phi(\rho/\pi,e),0,e)$ and $s(\phi(\rho/\pi,e),0,e)$ as the level of riskless and risky investments the naive agent would make if pension contributions were null.[25] Two cases may arise depending on the level of B:

If $B \leq b(\phi(\rho/\pi,e),0,e)$, agents top up the investment in the riskless asset, so that the additional investment in the riskless asset is $b^{EP} = b(\phi(\rho/\pi,e),0,e) - B$, with $\partial b^{EP}/\partial B = -1$. Since total investment ($B + b^{EP}$) in the riskless asset is not affected by B, and remains equal to the desired level of an agent with bias $\phi(\rho/\pi,e)$, the investment in the risky asset is not affected either, so that $\partial s^{EP}/\partial B = 0$.

Conversely, if $B > b(\phi(\rho/\pi,e),0,e)$, agents prefer a lower level of investment in the riskless asset. They therefore set $b^{EP} = 0$, so that $\partial b^{EP}/\partial B = 0$. The first order condition with respect

[25]Under Assumption 1, the solution to the individual problem in the absence of pension contributions is always interior.

to s^{EP} is

$$-1 + r_s \beta \phi(\rho/\pi, e) \pi u'(s^{EP} r_s + B r_b) = 0.$$

Fully differentiating this condition, we also find that

$$\frac{\partial s^{EP}}{\partial B} = -\frac{r_b}{r_s} < 0,$$

so that, like in the previous sections, s^{EP} decreases whenever B increases.

Let us now define the *ex ante* (EA) expected utility. Since agents are naive, they consider education as useless. The *ex ante* expected utility is then the perceived expected utility they *believe* they will maximize. It has the following expression:

$$U_{\rho}^{EA}(s, b) = w - s - b - pe - B - T + \beta \left[\rho u(s r_s + (b + B) r_b) + (1 - \rho) u((b + B) r_b) \right]. \quad (28)$$

where we recall that $\rho = \phi(\alpha, 0) = \alpha\pi$ is the probability of a high return as perceived by an agent with bias α. Maximizing the above function with respect to b and s, we obtain $b^{EA} \equiv b(\rho/\pi, B, e)$ and $s^{EA} \equiv s(\rho/\pi, B, e)$ which are the (*ex ante*) investment levels that the agents believe they will be making. They are defined by the following first-order conditions

$$\begin{aligned} &-1 + r_b \beta \left[\rho u'(d_1^{EA}) + (1 - \rho) u'(d_2^{EA}) \right] \leq 0, \\ &-1 + r_s \beta \rho u'(d_1^{EA}) \leq 0, \end{aligned} \quad (29)$$

with $d_1^{EA} = r_b(b^{EA} + B) + r_s s^{EA}$ and $d_2^{EA} = r_b(b^{EA} + B)$.

Let us now specify the incentive constraint of agents of type ρ. At the time of reporting, they do not take into account their *true* preferences. Instead, they maximise $U_{\rho}^{EA}(s, b)$ as defined by (28). The perceived ex-ante expected utility of a type ρ declaring to be of type $\tilde{\rho}$, is

$$\begin{aligned} U_{\tilde{\rho}}^{EA}(\tilde{s}, \tilde{b}) &= w - pe(\tilde{\rho}) - B - T(\tilde{\rho}) - \tilde{s} - \tilde{b} \\ &+ \beta \{ \rho u(r_s \tilde{s} + r_b(\tilde{b} + B)) + (1 - \rho) u(r_b(\tilde{b} + B)) \}, \end{aligned} \quad (30)$$

with $\tilde{s} = s(\rho/\pi, B, e(\tilde{\rho}))$ and $\tilde{b} = b(\rho/\pi, B, e(\tilde{\rho}))$. An incentive compatible public policy must then satisfy:

$$\frac{dU^{EA}_{\tilde{\rho}}(\tilde{s},\tilde{b})}{d\tilde{\rho}}\Big|_{\tilde{\rho}=\rho} = 0 \ \forall \rho,$$
$$\frac{d^2U^{EA}_{\tilde{\rho}}(\tilde{s},\tilde{b})}{d^2\tilde{\rho}}\Big|_{\tilde{\rho}=\rho} \leq 0 \ \forall \rho,$$

which are respectively the first-order and second-order local incentive compatibility constraints.

We show in Appendix 6 that the first-order incentive compatibility constraint simplifies to[26]

$$p\dot{e}(\rho) + \dot{T}(\rho) = 0. \tag{31}$$

To understand this condition, recall from Section 4.2, that if the government was offering the full-information policy, all (naive) agents would choose the allocation minimizing the first-period individual total contribution $T(\rho)+pe(\rho)$, without seeing the benefits from education on $\phi(\rho/\pi, e)$. Hence, the contract is incentive compatible as long as condition (31) is satisfied, that is as long as an increase (resp. decrease) in the level of education (which is costly to the agent) is compensated by a decrease (resp. increase) in the level of the lump-sum tax. This implies that independently from their type ρ, all agents obtain the same level of $T(\rho) + pe(\rho)$.[27]

5.2 The government's problem

Let us now derive the problem of the government given the informational constraints. The government seeks to maximize the expected *true* utility of agents with biases $\phi(\rho/\pi, e)$. This corresponds to the utility without misperception bias evaluated at the *ex post* investments b^{EP} and s^{EP} that agents choose after benefiting from education. This writes[28]

$$\begin{aligned} U^T_\rho(B, e, T) =& w - s^{EP} - b^{EP} - pe - B - T \\ &+ \beta\{\pi u(r_s s^{EP} + r_b(b^{EP} + B)) + (1-\pi)u(r_b(b^{EP} + B))\}, \end{aligned}$$

where b^{EP} and s^{EP} are defined, respectively, by (26) and (27).

[26] This is equivalent to a monotonicity constraint, which can replace second-order local incentive constraints. See Appendix 6 for further discussion.

[27] This violates the single-crossing condition.

[28] For ease of notation, we have dropped the arguments in $b^{EP}(\phi(\rho/\pi, e), B, e)$ and $s^{EP}(\phi(\rho/\pi, e), B, e)$.

We make a change of variable by denoting $z(\rho) = T(\rho) + pe(\rho)$, with $\dot{z}(\rho) = 0$ under the incentive compatibility constraint (31). Under asymmetric information, the problem of the government then writes as follows

$$\begin{aligned} & \max_{B,e(\rho),z} \int_{\pi\alpha_{min}}^{1} \Phi(U_\rho^T(B,e,T)) f_\rho(\rho) d\rho \\ = \quad & \int_{\pi\alpha_{min}}^{1} \Phi(w - z - B - s^{EP} - b^{EP} + \beta\{\pi u(r_s s^{EP} + r_b(b^{EP} + B)) + (1-\pi) u(r_b(b^{EP} + B))\}) f_\rho(\rho) d\rho \\ & \text{s.t.} \int_{\pi\alpha_{min}}^{1} [z - (p+q)e] f_\rho(\rho) d\rho \geq 0. \end{aligned}$$

Denoting by λ the Lagrange multiplier associated to the resource constraint, the FOCs are written as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial B} &= \int_{\pi\alpha_{min}}^{1} \Phi'(U_\rho^T(B,e,T))[\{-1 + \beta r_b \left[\pi u'(d_1^{EP}) + (1-\pi) u'(d_2^{EP})\right]\} \left(\frac{\partial b^{EP}}{\partial B} + 1\right) \\ & \quad + \{-1 + \beta \pi r_s u'(d_1^{EP})\} \frac{\partial s^{EP}}{\partial B}] f_\rho(\rho) d\rho = 0, \qquad (32) \\ \frac{\partial \mathcal{L}}{\partial e(\rho)} &= \Phi'(U_\rho^T(B,e,T))[\{-1 + \beta r_b \left[\pi u'(d_1^{EP}) + (1-\pi) u'(d_2^{EP})\right]\} \frac{\partial b^{EP}}{\partial e} \\ & \quad + \{-1 + \beta \pi r_s u'(d_1^{EP})\} \frac{\partial s^{EP}}{\partial e}] f_\rho(\rho) - \lambda f_\rho(\rho)(p+q) = 0 \quad \forall \rho \in [\pi\alpha_{min}, 1], \\ \frac{\partial \mathcal{L}}{\partial z} &= \int_{\pi\alpha_{min}}^{1} [-\Phi'(U_\rho^T(B,e,T)) + \lambda] f_\rho(\rho) d\rho = 0. \qquad (33) \end{aligned}$$

5.2.1 Optimal pension benefits

Let us first study the asymmetric information level of B. As shown above, if $B \leq b(\phi(\rho/\pi, e), 0, e)$, agents are not constrained in their investment choices *ex-post* and they will top up the investment in the riskless asset for an amount $b^{EP} = b(\phi(\rho/\pi, e), 0, e) - B$. In that case, $\partial b^{EP}/\partial B = -1$ and $\partial s^{EP}/\partial B = 0$ so that the RHS of (32) is always nil. Any level of $B \leq b(\phi(\rho/\pi, e), 0, e)$ yields the same level of welfare for agents of type ρ, which is equal to the laissez-faire one.

If $B > b(\phi(\rho/\pi, e), 0, e)$, we have $b^{EP} = 0$ so that $\partial b^{EP}/\partial B = 0$ and $\partial s^{EP}/\partial B = -r_b/r_s$. The RHS of (32) then simplifies to

$$\begin{aligned} & \int_{\pi\alpha_{min}}^{1} \Phi'(U_\rho^T(B,e,T)) f_\rho(\rho) [\{-1 + \beta r_b \left[\pi u'(d_1^{EP}) + (1-\pi) u'(d_2^{EP})\right]\} \\ & - \{-1 + \beta \pi r_s u'(d_1^{EP})\} \frac{r_b}{r_s}], \qquad (34) \end{aligned}$$

We need to study separately the case of pessimistic and optimistic agents. Consider first pessimistic agents, i.e. those with $\rho < \pi$. Using the individual FOCs with respect to b^{EP} and s^{EP}, i.e. equations (26) and (27) set to equality, we can deduce that

$$-1 + \beta r_b \left[\pi u'(d_1^{EP}) + (1-\pi)u'(d_2^{EP})\right] < 0,$$
$$-1 + \beta r_s \pi u'(d_1^{EP}) > 0.$$

Indeed, no matter the level of education, pessimistic agents would, ex post, invest too little in the risky asset and too much in the riskless one. This implies that (34) is always negative. Any pension contribution $B > b(\phi(\rho/\pi, e), 0, e)$ reduces the welfare of pessimistic agents with respect to the case where $B \in [0, b(\phi(\rho/\pi, e), 0, e)]$.

Consider now the case of optimistic agents, i.e. with $\rho > \pi$. If $B > b(\phi(\rho/\pi, e), 0, e)$, the RHS of (32) equivalent to (18) and, as we have shown in Section 4.1, it implies that $B = b^{FB}$ maximizes the welfare of all optimistic agents.

Hence, under asymmetric information, it is optimal to set $B = b^{FB}$ for every agent. This level of pension contribution forces optimistic agents to invest the right amount in the riskless asset. Any level of $B \leq b(\phi(\rho/\pi, e), 0, e)$ provides pessimistic agents with the same utility level, and with a higher utility than any $B > b(\phi(\rho/\pi, e), 0, e)$.[29]

5.2.2 Optimal financial education

Using (33), we can rearrange the FOC for $e(\rho)$ as follows:

$$\frac{\Phi'(U_\rho^T(B,e,T))}{\int_{\pi\alpha_{min}}^1 \Phi'(U_\rho^T(B,e,T)) f_\rho(\rho) d\rho} [\left\{-1 + \beta r_b \left[\pi u'(d_1^{EP}) + (1-\pi)u'(d_2^{EP})\right]\right\} \frac{\partial b^{EP}}{\partial e}$$
$$+ \left\{-1 + \beta \pi r_s u'(d_1^{EP})\right\} \frac{\partial s^{EP}}{\partial e}] = p + q, \qquad (35)$$

which can be compared with its full-information counterpart (22). The two expressions are the same except for the fraction on the LHS of the equality.

Differentiating $U_\rho^T(B,e,T)$ with respect to ρ and using the incentive compatibility condition,

[29] Note that $b^{FB} = b^{LF}(1) < b(\phi(\rho/\pi, e), 0, e)$ since $\phi(\rho/\pi, e) < 1$ for pessimistic agents.

we obtain that,

$$\frac{dU_\rho^T(B,e,T)}{d\rho} = \left\{-1+\beta r_b\left[\pi u'(d_1^{EP})+(1-\pi)u'(d_2^{EP})\right]\right\}\frac{\partial b^{EP}}{\partial\rho} + \left\{-1+\beta\pi r_s u'(d_1^{EP})\right\}\frac{\partial s^{EP}}{\partial\rho},$$

so that

$$\frac{dU^T}{d\rho} \geq 0 \text{ for } \rho \leq \pi,$$
$$\frac{dU^T}{d\rho} < 0 \text{ for } \rho > \pi.$$

Hence $\Phi'(U_\rho^T(B,e,T))$ is first decreasing in ρ until $\rho = \pi$ and then increasing up to 1, while $\int_{\pi\alpha_{min}}^1 \Phi'(U_\rho^T(B,e,T))f_\rho(\rho)d\rho = \lambda$ is constant. We can therefore define two thresholds of the misperception degrees, $\underline{\rho}$ and $\overline{\rho}$ with $\underline{\rho} < \pi < \overline{\rho}$ such that

$$\frac{\Phi'(U_{\underline{\rho}}^T(B,e,T))}{\int_{\pi\alpha_{min}}^1 \Phi'(U_\rho^T(B,e,T))f_\rho(\rho)d\rho} = \frac{\Phi'(U_{\overline{\rho}}^T(B,e,T)}{\int_{\pi\alpha_{min}}^1 \Phi'(U_\rho^T(B,e,T))f_\rho(\rho)d\rho} = 1.$$

For any level of $\rho < \underline{\rho}$ and $\rho > \overline{\rho}$, the ratio in (35) is greater than 1 so that education is distorted upward under asymmetric information. Conversely, for any $\underline{\rho} < \rho < \overline{\rho}$, this ratio is smaller than one, leading to a downward distortion in comparison to the full-information solution. Intuitively, the government wishes to redistribute towards severely biased agents. Under asymmetric information to avoid mimicking, education has to be distorted up for the severely biased agents (the mimickee) and down for the moderately biased agents (the mimickers) .

Let us now study how the level of education $e(\rho)$ under asymmetric information compares with its full information level. Figure 1 represents two possible cases depending on the comparison between $\underline{\rho}/\pi$ and $\overline{\rho}/\pi$ on the one hand, and α_1 and α_2, which correspond to the full information thresholds ensuring an interior solution for education (see Proposition 3), on the other hand. In turn, this comparison depends on the costs of education, the distribution of the bias, and the shape of the social welfare function.[30] In the figure, the black (resp. dashed) line represents the education level under symmetric (resp. asymmetric) information.

[30]In Figure 1, we have implicitly assumed that the distribution of $e(\rho)$ is symmetric, which is not necessarily the case. A similar reasoning applies to the non symmetric case.

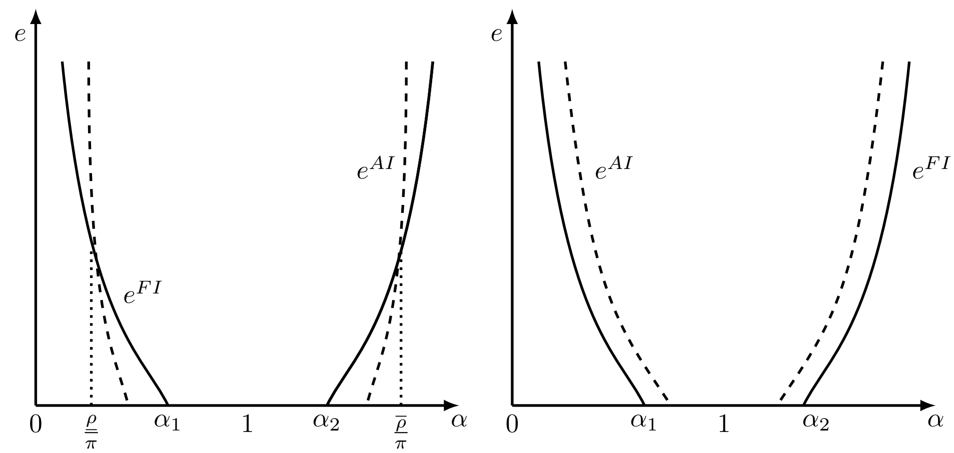

Figure 1: *Education as a function of the bias under the full (FI) and the asymmetric information (AI) solutions.*

On the left panel of Figure 1, only the education level of the very pessimistic and of the very optimistic agents are distorted upwards, while the education level of agents with an intermediate level of the bias is distorted downward. As in the full information case, it is optimal to provide no education to moderately biased agents, but the interval over which no education should be provided increases. This case may arise when the proportion of strongly biased agents is high or when their bias is very severe (i.e. for a fat-tailed, high-kurtosis distribution of α). Indeed, this corresponds to a case where the correction of their bias is relatively costly for two reasons. First, education entails a cost to the government (equal to $qe(\rho)$ per agent of type ρ). Second, in order to satisfy the incentive constraint, very biased agents need to be compensated for their high level of education (which costs them $pe(\rho)$) with a low tax (they may even receive a subsidy, $T(\rho) < 0$). Hence, providing a high level of education to these very biased agents puts a lot of financial pressure on the government. In that situation, it is then optimal to distort downward the education of the less-biased agents, in order to tax them more, which both relaxes the incentive compatibility constraints and generates public resources.

In the second case, a larger number of agents will get education under asymmetric information, and all levels of education will be weakly distorted upward. This case may arise when satisfying the monotonicity condition is not too demanding in terms of public resources, i.e. when the level of education $e(\rho)$ does not have to be compensated by very low levels of taxation through $T(\rho)$. This is more likely to happen when the proportion of strongly biased agents is low or the size of the biases is not too important. This is also the case if the public cost of education q is not too important.

As we showed above, both the public and the private costs of education play an important role. The public cost q plays a role in the government budget constraint, while p plays a role in the incentive compatibility constraint. For instance, if $p + q$ is low, in the full information case, most agents receive education, even those with a low bias. Distorting upward the level of education for all agents in the asymmetric information case is not too costly for the society, and this can be done through additional taxation. To the opposite, if $p + q$ is important, satisfying the incentive constraints is very costly to the society, and only the education of the severely biased agents can be distorted upward, while the education of moderately biased agents is distorted downward.

Our findings are summarized in the following proposition.

Proposition 4 *Consider an economy where agents invest both in a risky and a riskless asset, and differ in their levels of misperception α of the risk associated to the risky asset. If a paternalistic government cannot observe individual biases and can intervene through a pension system and personalized levels of financial education, the optimal second-best contract consists in*

i) a mandatory pension contribution, $B = b^{FB}$;

ii) a uniform total transfer, $\bar{z} = T(\rho) + pe(\rho)$ for all agents with different misperception bias, $\rho = \alpha\pi$;

iii) individualized levels of financial education $e(\rho)$. In comparison to the full information case, financial education is always distorted upward for very biased agents (either optimistic or pessimistic), while it may be distorted upward or downward for agents with mild (positive or negative) biases. Education levels should be equal to zero for agents with very low levels of *misperception.*

6 Conclusions

This paper studies the optimal design of public intervention when agents need to make financial plans for retirement. To do so, agents can invest both in a riskless and a risky asset, but they exhibit a misperception bias regarding financial saving returns. Misperception takes the form of either pessimism or optimism regarding the return obtained from the risky asset. The government intervenes through a mandatory pension system which provides uniform benefits (e.g. a first-pillar pension system). It also provides personalized education so as to correct for agents' pessimism or optimism.

We first show that the public pension system is an imperfect instrument to correct for the misperception bias. Pessimistic agents can always complement investments in both assets, so that their welfare is unchanged in comparison to the laissez faire. Optimistic agents end up overinvesting in the risky asset.

We then assume that the government can offer personalized levels of education in exchange of taxation. We show that, under perfect information on the misperception biases, the government should impose a uniform level of pension contributions equal to the first-best level of investment in the non-risky asset and education increasing in the degree of misperception (whether agents are optimistic or pessimistic). Under asymmetric information, we show that the level of education is always distorted upward for agents with important misperception biases (who either under- or over-estimate financial returns) and can be distorted upward or downward for agents with mild misperception biases. Whether we end up in one or the other situation depends on the public and private costs of education as well as on the shape of the distribution of the misperception biases in the economy.

Our model also points to the difficulty of designing optimal contracts under asymmetric information when agents who are biased are also naive in the sense that they do not realise the impact of financial education on their choices. As such, the discrepancy between the choices the agents *believe* they are making (entering the incentive constraint) and the ones they effectively make (entering the government's problem), affects the asymmetric-information problem in a non trivial way.

Finally, our model makes some simplifications and could be extended in several directions. One of them is certainly the assumption that all agents have the same income. Assuming otherwise would justify the introduction of a more elaborated pension system that would achieve redistribution from high-income toward low-income individuals. Yet, under asymmetric information on both dimensions (income and misperception), the asymmetric-information problem would become more difficult to solve. One way to overcome this problem could be to assume a correlation between income and misperception. This is on our research agenda.

Second, throughout the paper we have assumed that the pension contribution is invested exclusively in the riskless asset. However, one could think about other types of pension systems where the contributions are also invested in the risky asset. As long as agents observe the composition of the mandatory investment and can top up, such a system would not be able to correct perfectly for misperception biases. Optimistic agents would top up the contribution invested in the risky asset, while pessimistic agents would top up the one invested in the riskless asset. There would then still be a role for financial education, which we expect to be qualitatively similar to the one described in our paper.

All in all, we believe that our model is a first step in better understanding the interplay between pensions and financial education. As mentioned in the introduction, the literature on this subject is quite scarce. At a time where pension systems face important financial issues, and where governments ought to find solutions ensuring their viability, this paper shows how optimal financial education should be designed so as to increase welfare.

References

[1] Arrondel, L., M. Debbich and F. Savignac, 2015, Stockholding in France: the Role of Financial Literacy and Information, *Applied Economics Letters*, 22 (16), 1315–1319.

[2] Bernheim, B.D. and A. Rangel, 2007, Toward Choice-Theoretic Foundations for Behavioral Welfare Economics, *American Economic Review*, 97, 464–470.

[3] Bernheim B.D. and A. Rangel, 2009, Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics, *Quarterly Journal of Economics*, 124, 51–104.

[4] Browning M. and T. F. Crossley, 2001, The Life-Cycle Model of Consumption and Saving, *Journal of Economic Perspectives*, 15 (3), 3-22.

[5] Corsini, L. and L. Spataro, 2014, Optimal decisions on pension plans in the presence of information costs and financial literacy, *Journal of Public Economic Theory*, 17 (3), 383-414.

[6] Cremer, H., P. De Donder, D. Maldonado and P. Pestieau, 2007, Voting over type and generosity of a pension system when some individuals are myopic, *Journal of Public Economics*, 91, 2041- 2061.

[7] Cremer, H., P. De Donder, D. Maldonado and P. Pestieau, 2008, Designing an optimal linear pension scheme with forced savings and wage heterogeneity, *International Tax and Public Finance*, 15, 547-562.

[8] Cremer, H., P. De Donder, D. Maldonado and P. Pestieau, 2009, Non linear pension schemes with myopia, *Southern Economic Journal*, 76, 86-99.

[9] Diamond, P., and B. Koszegi, 2003, Quasi-hyperbolic discounting and retirement, *Journal of Public Economics*, 87, 1839-1872.

[10] Dutta, J., S. Kapur and J. Orszag, 2000, A portfolio approach to the optimal funding of pensions, *Economics Letters*, 69, 201-206.

[11] Fisher, I. 1930, The Theory of Interest. *New York: MacMillan.*

[12] Friedman, M., 1957, A Theory of the Consumption Function, *Princeton: Princeton University Press for NBER.*

[13] Guiso, L. and Sodini, P., 2013, Household Finance: An Emerging Field, *Handbook of the Economics of Finance*, chap. 21, 1397-1532..

[14] Hurd and Rohwedder, 2011, Stock Price Expectations and Stock Trading, Rand W-P 938.

[15] Hudomiet, P., G. Kezdi, and R., Willis, 2011, Stock market crash and expectations of american households, *Journal of Applied Econometrics*, 26, 393-415.

[16] Japelli T. and M. Padula, 2013, Investment in financial literacy and saving decisions, *Journal of Banking & Finance*, 37, 2779-2792.

[17] Japelli, T and M. Padula, 2015, Investment in financial literacy, social security,and portfolio choice, *Journal of Pension Economics & Finance*, 14 (4), 369-411.

[18] Kezdi, G. and R. Willis, 2009, Stock market expectations and portfolio choice of American households, working paper.

[19] Lalime T. and P-C. Michaud, 2014, Litteratie financière et préparation à la retraite au Québec et dans le reste du Canada, *L'Actualité économique*, 90 (1), 23-45.

[20] Lusardi A., O. Mitchell and V. Curto, 2010, Financial Literacy among the young, *The Journal of Consumer Affairs*, 44 (2), 358-380.

[21] Lusardi A., O. Mitchell and V. Curto, 2014, Financial Literacy and financial sophistication in the older population, *Journal of Pension Economics and Finance*, 13 (4), 347-366.

[22] Lusardi A. and O. Mitchell, 2007, Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education, *Business Economics.*

[23] Lusardi, A. and O. Mitchell, 2011, Financial Literacy and Retirement Planning in the United States, *Journal of Pension Economics and Finance*, 10(4), pp. 509-525.

[24] Lusardi, A. and O. Mitchell, 2014, The Economic Importance of Financial Literacy: Theory and Evidence, *Journal of the Economic Literature*, 52 (1), 5-44.

[25] Lusardi, A., P.-C. Michaud and O. Mitchell, 2017, Optimal Financial Knowledge and Wealth Inequality, *Journal of Political Economy,* 125 (2), 431-477.

[26] Malmendier U. and S. Nagel, 2011, depression babies: do macroeconomic experiences affect risk taking?, *The Quarterly Journal of Economics*, 126 (1), 373-416.

[27] Matsen, E., and Thogersen, O., 2004, Designing social security - A portfolio choice approach, *European Economic Review*, 48, 883-904.

[28] OECD, 2005, Improving Financial Literacy Analysis of Issues and Policies, *OECD Publishing.*

[29] Ramsey, F., 1928, A Mathematical Theory of Saving.*Economic Journal*, 38, 543-559.

[30] Samuelson, P. A., 1975, Optimum social security in a life-cycle growth model, *International Economic Review*, 16, 539-544.

[31] Schwartz S., 2010, Can Financial Education improve financial literacy and retirement planning, *IRPP Study*, 12.

[32] Spataro, L. and Corsini, L., 2017, Endogenous financial literacy, saving and stock market participation, *FinanzArchiv*, 73(2), 135-162.

[33] Van Rooij, M. C., Lusardi A. and R. J. Alessie, 2011, Financial literacy and stock market participation. *Journal of Financial Economics*, 101, 449-472.

[34] Van Rooj M., A. Lusardi and R. Alessie, 2012, Financial literacy, retirement planning and household wealth, *The Economic Journal*, 122, 449-478.

Appendix

Comparative statics - Laissez faire

Replacing (9) into (8), we obtain

$$\frac{\partial U}{\partial b} = -1 + \frac{r_b}{r_s} + r_b\beta(1-\alpha\pi)u'(d_2) \leq 0.$$

Fully differentiating this expression with respect to α_i we obtain

$$\frac{\partial b^{LF}}{\partial \alpha} = \frac{\pi u'(d_2)}{r_b(1-\alpha\pi)u''(d_2)} \leq 0.$$

As a consequence, $d_2^{LF} = r_b b^{LF}$ is decreasing in α. Similarly, fully differentiating (9) with respect to α yields

$$\frac{\partial s^{LF}}{\partial \alpha} = -\frac{u'(d_1)}{r_s\alpha u''(d_1)} - \frac{r_b}{r_s}\frac{\partial b^{LF}}{\partial \alpha} \geq 0.$$

As a consequence, the derivative of $d_1^{LF} = r_b b^{LF} + r_s s^{LF}$ with respect to α is equal to

$$\frac{\partial d_1^{LF}}{\partial \alpha} = r_s\frac{\partial s^{LF}}{\partial \alpha} - r_b\frac{\partial b^{LF}}{\partial \alpha} \geq 0.$$

Decentralization of the first-best optimum

The decentralized individual problem under such a tax scheme can therefore be written as follows:[31]

$$\begin{aligned}\max_{s,b} U^D(s,b) &= w - s - b - T + \beta[\alpha\pi u(s(1-\tau)r_s + b(1-\theta)r_b) \\ &+ (1-\alpha\pi)u(b(1-\theta)r_b)]\end{aligned}$$

where D stands for Decentralization. The FOCs are:

$$\frac{\partial U^D}{\partial b} = -1 + \beta r_b\left[\alpha\pi u'(d_1^D) + (1-\alpha\pi)u'(d_2^D)\right](1-\theta) \leq 0, \quad (36)$$

$$\frac{\partial U^D}{\partial s} = -1 + \beta\alpha\pi r_s u'(d_1^D)(1-\tau) \leq 0. \quad (37)$$

Comparing these FOCs with the first best ones (5) and (6), we obtain, after some rearrangements, that the tax levels decentralizing the first-best levels of investment in the risky and the riskless

[31] For ease of notation, we omit the argument α in the functions below.

assets are:

$$\tau^*(\alpha) = 1 - \frac{1}{\alpha},$$
$$\theta^*(\alpha) = 1 - [\frac{\pi u'(\bar{d}_1) + (1-\pi)u'(\bar{d}_2)}{\alpha\pi u'(\bar{d}_1) + (1-\alpha\pi)u'(\bar{d}_2)}].$$

Let us now briefly comment on the levels of the lump-sum taxes. The only source of heterogeneity relates to misperceptions in the probability that the risky asset yields a favorable return, and the taxes $\{\tau^*(\alpha), \theta^*(\alpha)\}$ enable the planner to obtain the optimal amounts of risky and riskless assets. These levels are uniform, so that at the decentralized equilibrium every agent invest $s^{LF}(1) = s^D(\alpha)(1-\tau^*(\alpha))$ in the risky asset and $b^{LF}(1) = b^D(\alpha)(1-\theta^*(\alpha))$ in the riskless asset, where $b^D(\alpha)$ and $s^D(\alpha)$ are the solutions to (36) and (37). Lump-sum taxes here are necessary to equalize expected utilities U^T (see equation 4), which, together with the fact that all agents invest $s^{LF}(1)$ and $b^{LF}(1)$, is achieved by ensuring an equal level of first-period consumption for all agents. The optimal levels of lump sum taxes are then given by eq. (13).

Financial Education: comparative statics with respect to e

Replacing (20) into (19), we have that

$$\frac{\partial U}{\partial b} = -1 + \frac{r_b}{r_s} + r_b\beta(1-\phi(\alpha,e)\pi)u'(d_2) \leq 0.$$

Fully differentiating this equation with respect to e, we get

$$\frac{\partial b}{\partial e} = \frac{\phi_e(\alpha,e)\pi u'(d_2)}{(1-\phi(\alpha,e)\pi)u''(d_2)r_b}.$$

Since $\phi_e(\alpha,e) = (1-\alpha)$, this derivative is negative if $\alpha < 1$, and positive otherwise. Fully differentiating (20) with respect to e, we obtain

$$\frac{\partial s}{\partial e} = -\frac{\phi_e(\alpha,e)u'(d_1)}{r_s\phi(\alpha,e)u''(d_1)} - \frac{r_b}{r_s}\frac{db}{de},$$

which is positive for $\alpha < 1$ and negative otherwise.

Proof of Proposition 3

Let us first establish some important comparative statics results. Replacing for (20) into (19), and differentiating it with respect to α, we obtain

$$\frac{\partial b}{\partial \alpha} = \frac{\phi_\alpha \pi u'(d_2)}{r_b(1-\phi\pi)u''(d_2)} \leq 0.$$

Similarly, differentiating (20) with respect to α yields

$$\frac{\partial s}{\partial \alpha} = -\frac{\phi_\alpha u'(d_1)}{r_s \phi u''(d_1)} - \frac{r_b}{r_s}\frac{\partial b}{\partial \alpha} \geq 0.$$

where we have dropped the arguments in $\phi(\alpha, e)$ for ease of notation.

Using these expressions we also find

$$\frac{\partial d_1}{\partial \alpha} = r_s \frac{\partial s}{\partial \alpha} + r_b \frac{\partial b}{\partial \alpha} = -\frac{\phi_\alpha u'(d_1)}{\phi u''(d_1)} \geq 0$$

and

$$\frac{\partial d_1}{\partial e} = r_s \frac{\partial s}{\partial e} + r_b \frac{\partial b}{\partial e} = -\frac{\phi_e u'(d_1)}{\phi u''(d_1)},$$

which is positive if $\alpha < 1$, and negative if $\alpha > 1$.

Furthermore,

$$\frac{\partial^2 b}{\partial \alpha \partial e} = \frac{\partial u'(d_2)/u''(d_2)}{\partial d_2} r_b \frac{\partial b}{\partial e} \frac{\phi_\alpha \pi}{r_b(1-\phi\pi)} + \frac{\pi u'(d_2)}{r_b u''(d_2)} \frac{\pi - 1}{(1-\phi\pi)^2}.$$

With CARA preferences, the first term is equal to zero, so that

$$\frac{\partial^2 b}{\partial \alpha \partial e} = \frac{\pi u'(d_2)}{r_b u''(d_2)} \frac{\pi - 1}{(1-\phi\pi)^2} \geq 0.$$

Similarly,

$$\frac{\partial^2 s}{\partial \alpha d e} = -\frac{\phi_\alpha}{\phi} \frac{\partial u'(d_1)/u''(d_1)}{\partial d_1} \frac{\partial d_1}{\partial e} + \frac{u'(d_1)}{r_s u''(d_1)} \frac{1}{\phi^2} - \frac{r_b}{r_s} \frac{\partial^2 b}{\partial \alpha \partial e}.$$

With CARA preferences, the first term is equal to zero, so that

$$\frac{\partial^2 s}{\partial \alpha \partial e} = \frac{u'(d_1)}{r_s u''(d_1)} \frac{1}{\phi^2} - \frac{r_b}{r_s} \frac{\partial^2 b}{\partial \alpha \partial e} \leq 0.$$

We can now evaluate the sign of $\partial e/\partial \alpha$. Differentiating (22) with respect to α and using the implicit function theorem, we obtain that $de/d\alpha$ has the same sign as

$$\begin{aligned}
\frac{\partial^2 W}{\partial e \partial \alpha} &= \frac{\partial d_1}{\partial \alpha}\left(\frac{\partial s}{\partial e} r_s + \frac{\partial b}{\partial e}\beta r_b\right)\beta\pi u''(d_1) + \beta r_b^2 (1-\pi)\frac{\partial b}{\partial \alpha}\frac{\partial b}{\partial e} u''(d_2) \\
&\quad + \left[-1 + \beta r_b \left(\pi u'(d_1) + (1-\pi)u'(d_2)\right)\right] \frac{\partial^2 b}{\partial \alpha \partial e} \\
&\quad + \left[-1 + \beta r_s \pi u'(d_1)\right] \frac{\partial^2 s}{\partial \alpha \partial e} \\
&= \frac{\partial d_1}{\partial \alpha}\frac{\partial d_1}{\partial e}\beta\pi u''(d_1) + \beta r_b^2 (1-\pi)\frac{\partial b}{\partial \alpha}\frac{\partial b}{\partial e} u''(d_2) \\
&\quad + \left[-1 + \beta r_b \left(\pi u'(d_1) + (1-\pi)u'(d_2)\right)\right] \frac{\partial^2 b}{\partial \alpha \partial e} \\
&\quad + \left[-1 + \beta r_s \pi u'(d_1)\right] \frac{\partial^2 s}{\partial \alpha \partial e}.
\end{aligned}$$

According to the comparative statics derived above, this expression is negative for $\alpha < 1$ and positive for $\alpha > 1$.

Informational constraints

Using the reported utility of a type-ρ agent, expression (30), we derive the first-order and second-order local incentive constraints. This yields

$$\frac{dU^{EA}_{\tilde{\rho}}(\tilde{s},\tilde{b})}{d\tilde{\rho}}\Big|_{\tilde{\rho}=\rho} = -p\dot{e}(\tilde{\rho}) - \dot{T}(\tilde{\rho}) + \dot{\tilde{s}}[-1+\beta\rho r_s u'(\tilde{d}_1)] + \dot{\tilde{b}}[-1+\beta r_b(\rho u'(\tilde{d}_1)+(1-\rho)u'(\tilde{d}_2))] = 0, \tag{38}$$

$$\frac{d^2U^{EA}_{\tilde{\rho}}(\tilde{s},\tilde{b})}{d^2\tilde{\rho}}\Big|_{\tilde{\rho}=\rho} = -p\ddot{e}(\tilde{\rho}) - \ddot{T}(\tilde{\rho}) + \ddot{\tilde{s}}[-1+\beta\rho r_s u'(\tilde{d}_1)] + \ddot{\tilde{b}}[-1+\beta r_b(\rho u'(\tilde{d}_1)+(1-\rho)u'(\tilde{d}_2))] + \beta\rho u''(\tilde{d}_1)r_s^2\dot{\tilde{s}}^2 + \beta r_b^2(\rho u''(\tilde{d}_1)+(1-\rho)u''(\tilde{d}_2))\dot{\tilde{b}}^2 \leq 0. \tag{39}$$

Let us now obtain the monotonicity constraint. Evaluating $\tilde{\rho} = \rho$, we can rewrite equation (38) as follows:

$$-p\dot{e}(\rho) - \dot{T}(\rho) + \dot{s}(\rho)[-1+\beta\rho r_s u'(d_1)] + \dot{b}(\rho)[-1+\beta r_b(\rho u'(d_1)+(1-\rho)u'(d_2))] = 0. \tag{40}$$

Two cases are possible. If B is not binding *ex ante*, i.e. $B \leq b(\rho/\pi, B, e)$, b and s are interior and (**??**) and (29) hold with equality. In this case, the second and third terms on the LHS of (40) are equal to zero, so that the expression simplifies to $p\dot{e}(\rho) + \dot{T}(\rho) = 0$.

In the same way, if B is binding *ex ante*, i.e. $B > b(\rho/\pi, B, e)$, we have from the individual problem, that $b(\rho/\pi, B, e) = 0$ and $s(\rho/\pi, B, e)$ is interior, so that $\dot{b} = 0$ and (29) holds with equality. The above condition again simplifies to $p\dot{e}(\rho) + \dot{T}(\rho) = 0$.

This condition also implies that the second-order condition in (39) is satisfied in $\rho = \widetilde{\rho}$. The first two terms in (39) are equal to zero as long as $p\dot{e}(\rho) + \dot{T}(\rho) = 0$, while the second and third terms are equal to zero using the individual first-order conditions with respect to b and s. The only remaining terms are the last two, which are unambiguously negative.