

Optimal fund menus

Jakša Cvitanić¹ | Julien Hugonnier²

¹ Division of the Humanities and Social Sciences, Caltech, Pasadena, CA, USA

² EPFL, Swiss Finance Institute, and CEPR, Quartier UNIL Chamberonne, Lausanne, Switzerland

Correspondence

Jakša Cvitanić, Division of the Humanities and Social Sciences, Caltech. M/C 228–77, 1200 East California Blvd., Pasadena CA 91125, USA.

Email: cvitanic@hss.caltech.edu

Funding information

NSF, Division of Mathematical Sciences, Grant/Award Number: 1810807.

Abstract

We study the optimal design of a menu of funds by a manager who is required to use linear pricing and does not observe the beliefs of investors regarding one of the risky assets. The optimal menu involves bundling of assets and can be constructed from the solution to a calculus of variations problem that optimizes over the indirect utility that each type receives. We provide a complete characterization of the optimal menu and show that the need to maintain incentive compatibility leads the manager to offer funds that are inefficiently tilted towards the asset that is not subject to the information friction.

KEYWORDS

asset bundling, linear pricing, mutual fund menus, screening

JEL CLASSIFICATION

C62, C71, D42, D82, G11

1 | INTRODUCTION

One of the salient characteristics of the mutual fund industry is the proliferation of products that results from the fact that each investment firm offers a large number of funds that often overlap significantly.¹ In this paper, we propose a novel, information-based theory of mutual fund families that explains this proliferation. Our reasoning is simple. Consider a firm that provides investment services to a population of heterogeneous investors and assume that the manager of the firm knows the distribution of investors' characteristics but does not observe the individual type of each investor and thus faces an adverse selection problem. In such a setting, the manager needs to design its offering to screen investors and we claim that under a linear pricing constraint, the optimal strategy is to offer a menu of combinations of the risky assets – that is, funds – constructed to be differentially attractive to different types of investors.

To illustrate this mechanism, we consider a static model with one riskless asset and two risky assets. The market is populated by a risk neutral investment firm/manager and a continuum of mean-variance investors who can only access the risky assets through the firm. Investors agree on the variances of returns as well as on the expected return of the first asset but differ in their beliefs regarding the expected return on the second asset. The interpretation of this assumption is that the first risky asset represents a *familiar* investment vehicle such a broad domestic index about which information is freely available and widespread, whereas the second risky asset represents a less familiar asset such as a foreign index about which information is less easily gathered.

The manager of the firm knows the distribution of investor beliefs, but does not observe the type of each individual investor. As in a screening model with multiple goods (see, e.g., Wilson (1993) and Armstrong (1996)), the manager is allowed to offer combinations of assets and a corresponding pricing scheme, but we depart from the canonical screening setting in two important ways to take into account the specificities of the mutual fund market. First, we allow investors to combine the funds offered by the manager subject to a no-short-selling constraint. Second, we follow the US regulation of investment advisors by requiring the manager to use a linear pricing rule that specifies the fees on each fund as a fraction of assets under management.² Absent this constraint the well-known screening solution applies and the optimal strategy is to let investors trade the risky assets separately subject to a linear fee for the familiar asset and a nonlinear pricing scheme with quantity discounts for the non-familiar asset.³ On the contrary, if pricing is required to be linear then the only way to screen investors is to bundle assets into funds that each deliver a specific exposure to the risk factors, and one of our main contributions is to show how to construct the optimal fund menu.

The solution method we develop consists of three steps. First, we establish that a version of the revelation principle holds in our model. This allows us to restrict the manager to menus in which funds are indexed by investor types and which have the property that each investor finds it optimal to invest only in the fund targeted to his type. In the second step, we show that this incentive compatibility constraint can be reduced to a family of differential inequalities and use this formulation to establish that the optimal menu can be characterized in terms of the solution to a constrained calculus of variations problem. In the third and last step, we provide a complete analysis of the Euler–Lagrange equations associated with this problem and use the unique solution to these equations to explicitly construct the optimal menu.

The analysis of the optimal menu allows us to study the combined effect of the information and pricing frictions at play in our model. Consider first the impact of the linear pricing constraint when taking as given the information friction. We show that linear pricing reduces the amount of fees collected by the manager, increases the participation of investors as well as their aggregate welfare, and even results in strict Pareto improvements for all investors if the information friction is not too intense. Therefore, our results provide a justification for regulations, such as the 1970 Amendment to the Investment Advisors Act, that restrict the price setting ability of investment firms. Imposing linear pricing prevents the manager from using prices to discriminate among investors and instead leads him to rely on bundling as a screening device. Specifically, we show that the familiar asset is part of the menu – because it is not affected by the information friction – but that it is never optimal to offer the two assets separately. Our findings therefore contribute to the literature on asset bundling (see, e.g., Adams and Yellen (1976), Spence (1980), and McAfee et al. (1989)) by providing conditions under which linear pricing makes mixed bundling optimal.

Consider next the impact of the information friction taking the linear pricing constraint as given. We show that given linear pricing and complete information, it is also optimal to offer a

menu of funds. Comparing this menu to the optimal menu of the asymmetric information case allows to elicit the effects of the information friction. In particular, we show that the need to maintain incentive compatibility leads the manager to propose funds that are more tilted towards the familiar asset than those he would have offered under complete information. Our theory of fund families thus provides an alternative, information-based explanation for the well-known home bias according to which investors tend to over-invest in domestic/familiar assets.⁴ To provide intuition for this result, we show that, if the full information menu was offered in the asymmetric information case, investors would have an incentive to underreport their beliefs to benefit from the better conditions offered to more pessimistic investors. To prevent this from happening the manager needs to make the funds that target more pessimistic investors less attractive to more optimistic investors, and this is achieved by increasing the share of the familiar asset in all the funds.

Our base case includes a single investment firm that faces a population of investors, as would be the case when considering the provision of retirement accounts to the employees of a company. To introduce a form of competition, we study an extension in which investors can also access the familiar asset through an outside fund at some exogenous fee rate. We show that three cases may occur. If the exogenous fee rate is higher than the fee rate the manager would have offered for the familiar asset absent competition, then the outside fund is dominated. If the outside rate is lower than the optimal rate but still sufficiently high then competition leads the manager to exclude a fringe of pessimistic investors from the non-familiar asset market. Despite this exclusion, all investors benefit from the presence of the outside fund because its lower fee rate more than compensates for the lack of exposure to the non-familiar asset. As the outside fee rate decreases, investors become less willing to acquire exposure to the familiar asset otherwise than through the outside fund. This makes it harder for the manager to screen by bundling, and we show that there is a threshold below which the optimal strategy is to unbundle the assets. In this case, the optimal menu still excludes a fringe of pessimistic investors but can be implemented by offering the familiar asset at the market fee rate and the non-familiar asset at a constant fee rate that we determine in closed form.

While the tractability of our model rests on stark assumptions, we believe that the qualitative message of our paper is likely to remain valid in other settings. Instead of differing in their beliefs, investors could well differ along one or more other important dimension such as risk aversion, initial endowments, the assets they are willing to hold, or the risks they are exposed to. Furthermore, preferences need not be quadratic and there may exist more than two risky assets. These important extensions make the model less tractable because they lead to multidimensional screening problems (see, e.g., Rochet and Choné (1998)) but we believe that our solution method and the mechanism we highlight would play a important role in the solution. In particular, it is likely that linear pricing would still lead to asset bundling as a screening device.

Our paper relates to a large theoretical literature on delegated portfolio management. Hugonnier and Kaniel (2010) study a model close to ours but in which the fund manager faces a single investor about whom he has full information. Breton et al. (2010) extend the model of Hugonnier and Kaniel (2010) to the case where two managers compete and show that competition does not benefit investors because, in equilibrium, the funds offered by the two managers are colinear. In our model, we take as given that the pricing of funds must be linear. By contrast, Admati and Pfleiderer (1997), Carpenter (2000), Das and Sundaram (2002), Basak et al. (2007), Cuoco and Kaniel (2011), and Basak and Pavlova (2013) study the effects of different exogenous fee structures on allocations, social welfare, risk-taking, market efficiency, and asset prices, while Bhattacharya and Pfleiderer (1985), Ou-Yang (2003), Dybvig et al. (2010), and Cvitanic and Xing (2018) among

others, adopt an optimal contracting perspective in which investors control the compensation of the fund manager. Our model focuses on the design of an optimal fund menu in a static setting where the customer base is fixed. Therefore, it abstracts from some important dynamic considerations such as learning about managerial skill and its implications for the relation between past performance and fund flows. Examples of papers that examine the impact of investors' learning about managerial skills and/or technology include Lynch and Musto (2003), Dangi et al. (2008), Carlin and Manso (2011), Pastor and Stambaugh (2012), and Brown and Wu (2016) among others.

There are a few papers that model fund families. Mamaysky and Spiegel (2002) propose a model that explains the existence of many different fund families, while we focus on why there are many funds inside one family. In their model, each investment company gathers information that is specific and offers portfolios aimed at the subset of the population to which that information is most useful. In our model, the funds inside a given family adapt to the beliefs among its population of investors, but our finding is not necessarily at odds with those of Mamaysky and Spiegel (2002). In particular, they empirically document that when an investment firm introduces a new fund, it typically uses a strategy that places this fund in a different Morningstar category than its existing ones, which is in line with the fact that in our model, a new fund would only be introduced following a change in the customer base. Our findings are also in agreement with Gruber (1996), Khorana and Servaes (2012), and Massa (2000) who show both empirically and theoretically that product differentiation is an effective strategy for investment firms to maximize revenues. In recent work, Brown and Wu (2016) follows the approach of Berk and Green (2004) to develop a continuous-time model in which the performance of the funds offered by a sponsor carries information about the common skills and resources shared across the whole family, while Berk et al. (2017) propose a model of an investment firm that allocates its investors' capital to a population of heterogeneous fund managers who can each add value to the firm subject to decreasing returns to scale.

Our paper also contributes to the industrial organization literature on screening and asset bundling, see for example Adams and Yellen (1976), Spence (1980), McAfee et al. (1989), Wilson (1993), Armstrong (1996), and Stole (2001) among others. In particular, our paper can be seen as multiple goods extension of the model of Mussa and Rosen (1978) in which the monopolist is required to use linear pricing. Because of this constraint, the monopolist cannot resort to non-linear pricing as a mean of discriminating among his customers. Instead, she will use product bundling and our contribution is to show how the optimal menu of linearly priced bundles can be constructed. In a related contribution Rothschild (2015) also considers a screening problem with linear pricing, but his graphical analysis is limited to qualitative properties of the optimum. To the best of our knowledge, this paper is the first to analytically derive a solution to a screening problem with multiple goods and a linear pricing constraint.

The remainder of the paper is organized as follows. In Section 2, we present the model. In Section 3, we show how the design of an optimal fund menu can be reduced to the solution of a calculus of variations problem and provide a complete description of the optimal fund menu. In Section 4, we analyze the most salient properties of the optimal menu. Finally, in Section 5, we extend the base case model to allow investors a direct access to an outside fund that offers the familiar asset. Section 6 concludes. Appendix A derives the solution to our model in three important benchmark cases: the frictionless case where investors can freely access all assets, the asymmetric information case where the manager is allowed to use any pricing scheme, and the full information case where he is required to use linear pricing. Appendices B and C gather all the proofs.

2 | THE MODEL

We consider a static model of a financial market that consists in three assets: a riskless asset with gross return r and two risky assets whose *excess* returns are given by a vector $\varepsilon \in \mathbb{R}^2$ of independent random variables with unit variances. We interpret the first risky asset as representing a widely familiar asset, such as a broad domestic market index, about which investors have homogeneous beliefs and the second one as being a less familiar asset about which investors have dispersed beliefs.

The market is populated by a single risk-neutral investment manager and a unit measure of risk-averse investors. All market participants agree that returns have unit variances and that the expected gross excess return on the familiar asset, or equivalently its risk premium, is given by ξ for some constant $\xi > 0$, but investors differ in their beliefs regarding the expected return on the other asset. Specifically, we assume that each investor is associated with a type $\theta \in \Theta := [0, \theta_H]$ that represents her perception of the expected gross excess return on the non-familiar asset. Each investor knows her own type but the only information available to the manager is that the investors' types are uniformly distributed over Θ .

Investors have initial wealth w_0 and mean-variance preferences over terminal wealth. Specifically, we assume that the utility that an investor of type $\theta \in \Theta$ derives from terminal wealth w_1 is given by

$$u(\theta, w_1) := a(E_\theta[w_1] - rw_0) - \frac{a^2}{2}\text{var}_\theta[w_1], \quad (1)$$

where $a > 0$ captures the investors' risk-aversion and the subscript indicates that the computation of the mean and variance is performed under the probability measure P_θ associated with the investor's beliefs.⁵ Investors can trade the riskless asset but can only access the risky assets through the manager.⁶ In line with the regulation of investment advisors, we assume that the manager can only use *linear price schedules* that charge investors a constant fraction of the initial investment. Accordingly, a fund is specified by a pair (γ, ϕ) where $\gamma \in \mathbb{R}_+$ is the fee that the manager collects at the terminal time per dollar invested in the fund and $\phi \in \mathbb{R}^2$ represents the amounts invested in the two risky assets per dollar of assets under management.⁷ When offered, each such pair gives rise to a composite asset that investors can allocate capital to and which provides an *excess* return given by

$$\mathcal{R}(\gamma, \phi) := \phi^\top \varepsilon - \gamma \equiv \sum_{i=1}^2 \phi_i \varepsilon_i - \gamma. \quad (2)$$

Because the pricing of funds is constrained to be linear, the manager cannot rely on quantity discounts to screen investors as he would in the standard model of monopoly pricing under asymmetric information.⁸ Instead, he will exploit the fact that investors have different preferences for the risky assets by offering a menu of linearly priced funds that represent different combinations of exposures to these assets.

Definition 2.1. A fund menu is a collection $\mathbf{m} = (\gamma, \phi, \mathcal{M})$ where \mathcal{M} is a set that indexes the funds and $(\gamma, \phi) : \mathcal{M} \rightarrow \mathbb{R}_+ \times \mathbb{R}^2$ are functions that represent the fee rate and the vector of loadings of the funds on the risky assets.

Let \mathbf{m} be a given menu. In addition to linear pricing, another key feature of our model compared to a standard screening environment is that investors are not constrained to pick a single fund and can in fact combine funds to achieve their preferred exposure subject to a no short selling constraint. We capture this *non exclusivity* by taking the space $\mu_+(\mathcal{M})$ of nonnegative measures on \mathcal{M} as the action set of investors. If the investor allocates capital to the funds according to some measure q then

$$w_0 - \int_{\mathcal{M}} q(dm) = w_0 - q(\mathcal{M}), \tag{3}$$

is invested in the riskless and the induced terminal wealth is given by

$$\begin{aligned} w_1 = w_1(q, \mathbf{m}) &:= (w_0 - q(\mathcal{M}))r + \int_{\mathcal{M}} (r + \mathcal{R}(\gamma(m), \phi(m)))q(dm) \\ &= rw_0 + \int_{\mathcal{M}} \mathcal{R}(\gamma(m), \phi(m))q(dm). \end{aligned} \tag{4}$$

The optimization problem of an investor of type $\theta \in \Theta$ who takes the fund menu \mathbf{m} as given is then defined by

$$v_i(\theta, \mathbf{m}) := \sup_{q \in \mu_+(\mathcal{M})} u(\theta, w_1(q, \mathbf{m})), \tag{5}$$

and the aggregation of individual portfolio decisions generates a total amount of management fees given by

$$v_m(\mathbf{m}) := \frac{1}{\theta_H} \int_{\Theta \times \mathcal{M}} \gamma(m)q^*(dm; \theta, \mathbf{m}) d\theta, \tag{6}$$

where the measure $q^*(\cdot; \theta, \mathbf{m})$ is the best response of an investor of type $\theta \in \Theta$ to the menu \mathbf{m} . In accordance with the above definitions, a menu \mathbf{m}^* is optimal if it maximizes the total amount of fees in Equation (6).

Remark 2.2 (Fund composition and leverage). We do not require the fund loadings ϕ_1 and ϕ_2 to be positive or to sum up to one. As a result, the funds offered by the manager may in principle include short risky asset positions as well as long or short positions in the riskless asset. We show below that this assumption is without loss of generality as long as asset returns are independent. In particular, the offered funds are never short in any of the risky assets and, since linearly priced funds are defined up to constant, there always exist an *all-equity* implementation of the optimal menu in which none of the funds invest in the riskless asset.

Remark 2.3 (Investor leverage). Investors are not allowed to short the funds offered by the manager but we do not impose any constraint on the total amount $q(\mathcal{M})$ that each investor allocates to the funds. In particular, investors in our model can borrow at the risk-free rate if their preferred portfolio is such that $q(\mathcal{M}) \geq w_0$. This assumption may seem counterfactual but is without loss of generality if leverred funds are allowed since any borrowing can then be done through the funds. Even if the funds cannot use leverage, this assumption still

remains without loss of generality as long as the investors' initial wealth w_0 and/or risk aversion a is large enough since our preference specification implies that the capital optimality invested in risky assets is independent of initial wealth and inversely proportional to risk aversion.

Remark 2.4 (Specification of returns). The assumption that returns are independent and have unit variance simplifies the presentation of the results but does not entail any loss in generality as long as leverred funds are allowed. Indeed, a direct calculation shows that in our model investing through a fund (γ, ψ) in two correlated risky assets with non-unit variances is equivalent to investing in two uncorrelated risky assets with unit variances through the adjusted fund defined by

$$(\gamma, \phi) = \left(\gamma, \begin{bmatrix} \sigma_1 & \rho\sigma_2 \\ 0 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix} \psi \right), \quad (7)$$

where the constants $\sigma_1, \sigma_2 > 0$ and $\rho \in [-1, 1]$ denote, respectively, the standard deviations of asset returns and their correlation coefficient.

Remark 2.5 (Performance fees). Under specific conditions, the 1970 Amendment to the Investment Advisors Act of 1940 also allows the use of performance fees provided that they are symmetric around a benchmark. In our model, the imposition of a symmetric performance fee on a fund $\phi(m)$ would correspondant to charging

$$\gamma(m) + \delta(m)(\phi(m) - b(m))^\top \varepsilon, \quad (8)$$

per dollar initially invested in the fund where $b(m)$ represents a benchmark portfolio and $(\gamma(m), \delta(m))$ are nonnegative constants. Our main model does not include the possibility of such fees but this restriction is without loss of generality. Indeed, we show in Appendix B.4 that in our setting it is not optimal for the manager to use such fees.

3 | SOLUTION

3.1 | The revelation principle

Let $\xi(\theta) := (\xi, \theta)^\top$ be the vector of expected gross excess returns (i.e., risk premia) perceived by an investor of type $\theta \in \Theta$ and denote by

$$\pi(\theta, \phi) := \operatorname{argmax}_{q \in \mathbb{R}_+} u(\theta, rw_0 + q\mathcal{R}(1, \phi)) = \frac{1}{a\|\phi\|^2} (\phi^\top \xi(\theta) - 1)_+, \quad (9)$$

the amount that this investor would optimally invest in the fund $(1, \phi)$ when allocating his wealth between the riskless asset and that fund. Our first result is a version of the revelation principle, which shows that the manager can restrict her attention to the set of menus such that funds are indexed by types, fee rates are normalized to one, and each investor finds it optimal to only invest in the fund targeted to his type.

Proposition 3.1. *Given any menu $\bar{\mathbf{m}}$, there exists a menu $\mathbf{m} = (\gamma, \phi, \mathcal{M})$ such that:*

- (1) $\mathcal{M} = \Theta$;
- (2) $\gamma(\theta) = 1$ for all $\theta \in \Theta$;
- (3) $q^*(\sigma; \theta, \mathbf{m}) = \mathbf{1}_{\{\theta \in \sigma\}} \pi(\theta, \phi(\theta))$ for all $\theta \in \Theta$ and $\sigma \subseteq \Theta$.
- (4) $v_m(\mathbf{m}) = v_m(\bar{\mathbf{m}})$;
- (5) $v_i(\theta, \mathbf{m}) = v_i(\theta, \bar{\mathbf{m}})$ for all $\theta \in \Theta$.

Property 3 is the analogue in our setting of the *incentive compatibility constraint* in classical screening problems. Properties 4 and 5 mean that the manager and the investors are indifferent between the original menu and the new menu, which satisfies properties 1–3. The normalization of the fee is without loss of generality because funds are only defined up to a multiplicative constant. In the context of our model, this normalization serves two purposes: it reduces the choice of the manager to that of a *fund loading function* $\phi : \Theta \rightarrow \mathbb{R}^2$ and allows to easily compare the funds in a given menu by comparing the risk exposures that they offer.

Our next result provides a variational characterization of incentive compatibility that will be instrumental in our construction of the fund menu.

Proposition 3.2. *A loading function $\phi : \Theta \rightarrow \mathbb{R}^2$ is incentive compatible if and only if*

$$\phi(\theta')^\top \xi(\theta) - 1 - \frac{\phi(\theta)^\top \phi(\theta')}{\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1)_+ \leq 0, \quad (10)$$

for all $(\theta, \theta') \in \Theta^2$.

To understand the result, assume that the manager offers a loading function ϕ . If an investor of type $\theta \in \Theta$ perceives that the risk premium

$$E_\theta[\mathcal{R}(1, \phi(\theta))] = \phi(\theta)^\top \xi(\theta) - 1, \quad (11)$$

on the fund targeted to him is nonpositive, then Equation (10) requires that this investor also perceives all the other funds in the menu as offering negative risk premia and, thus, finds it optimal to only invest in the riskless asset. On the other hand, if the investor perceives that the risk premium on the fund targeted to him is strictly positive, then Equation (10) requires that, for any $\theta' \in \Theta$, the alpha

$$\begin{aligned} \alpha(\theta, \theta') &:= E_\theta[\mathcal{R}(1, \phi(\theta'))] - \frac{\text{cov}_\theta[\mathcal{R}(1, \phi(\theta)), \mathcal{R}(1, \phi(\theta'))]}{\text{var}_\theta[\mathcal{R}(1, \phi(\theta))]} E_\theta[\mathcal{R}(1, \phi(\theta))] \\ &= (\phi(\theta')^\top \xi(\theta) - 1) - \frac{\phi(\theta)^\top \phi(\theta')}{\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1), \end{aligned} \quad (12)$$

of fund $(1, \phi(\theta'))$ relative to fund $(1, \phi(\theta))$ be negative, so that including any other fund in his portfolio does not improve his risk-adjusted performance.

To elicit the nature of the information friction, it is useful to briefly consider the *first best* case in which the manager knows the type of each investor, but is still required to use linear pricing.

In this case, we show in Appendix A.3 that it is optimal to offer investors of type $\theta \in \Theta$ a single fund with loading function

$$\phi^\circ(\theta) = \frac{2\xi(\theta)}{\|\xi(\theta)\|^2}, \quad (13)$$

and fee rate equal to one. Substituting this loading function into the incentive compatibility condition (10) shows that

$$\phi^\circ(\theta')^\top \xi(\theta) - 1 - \frac{\phi^\circ(\theta)^\top \phi^\circ(\theta')}{\|\phi^\circ(\theta)\|^2} (\phi^\circ(\theta)^\top \xi(\theta) - 1) = \frac{\theta'(\theta - \theta')}{\|\xi(\theta')\|^2}, \quad (14)$$

is nonnegative for any pair $(\theta, \theta') \in \Theta^2$ such that $\theta' \leq \theta$. This shows that the first best fund menu is not incentive compatible and reveals that the adverse selection problem facing the manager is that, when offered the first best menu, any given investor has an incentive to pretend to be less optimistic than he really is.

3.2 | The relaxed problem

Propositions 3.1 and 3.2 imply that the manager's optimization problem reduces to the maximization of the integral

$$I(\phi) := (1/\theta_H) \int_{\Theta} \pi(\theta, \phi(\theta)) d\theta, \quad (15)$$

over the set Φ_0 of fund loading functions that satisfy Equation (10). To solve this problem, we further restrict the manager's choice set by imposing the technical requirement that the fund loading function belongs to the intersection $\Phi := \Phi_0 \cap AC(\Theta; \mathbb{R}^2)$ of Φ_0 with the space of absolutely continuous functions on Θ with values in \mathbb{R}^2 . The optimization problem that we consider is therefore given by

$$M := \sup_{\phi \in \Phi} I(\phi). \quad (\mathcal{P})$$

The main difficulty in solving this problem arises from the fact that Equation (10) cannot be dealt with using standard techniques because it involves the values of the unknown vector valued function at all points of the type space. To overcome this difficulty, we follow the *first-order approach* (see, e.g., Mirrlees (1971), Rochet (1987), and Rochet and Choné (1998)), which exploits the first-order condition induced by the incentive compatibility constraint (10) to show that, instead of optimizing over loading functions, the manager can optimize over the indirect utility

$$v(\theta) := u(\theta, rw_0 + \pi(\theta, \phi(\theta))\mathcal{R}(1, \phi(\theta))) = \frac{1}{2} \left(\frac{\phi(\theta)^\top \xi(\theta) - 1}{\|\phi(\theta)\|} \right)_+^2, \quad (16)$$

and marginal utility $\dot{v}(\theta)$ that her menu of funds delivers to each type of investor. Our first result in this direction relates incentive compatible loading functions to the indirect utility functions they induce.

Lemma 3.3. Assume that $\phi \in \Phi$ is incentive compatible. Then the indirect utility function defined in Equation (16) belongs to the space $AC(\Theta; \mathbb{R})$. Furthermore,

$$2v(\theta) \geq [\dot{v}(\theta)]^2, \quad (17)$$

and the corresponding optimal investment can be expressed as

$$\pi(\theta, \phi(\theta)) = F(\theta, v(\theta), \dot{v}(\theta)) := \frac{1}{a} \left(\theta \dot{v}(\theta) - 2v(\theta) + \xi \sqrt{2v(\theta) - [\dot{v}(\theta)]^2} \right), \quad (18)$$

for almost every $\theta \in \Theta$.

Let $\phi \in \Phi$ be an incentive compatible fund loading function. Relying on the above lemma, we have that the total amount of fees generated by the investors' best responses to the corresponding menu is given by

$$\theta_H I(\phi) = \int_{\Theta} \pi(\theta, \phi(\theta)) d\theta = \int_{\Theta} F(\theta, v(\theta), \dot{v}(\theta)) d\theta, \quad (19)$$

where v is the indirect utility function associated to ϕ through Equation (18). It follows that the manager's value function satisfies

$$\theta_H M = \sup_{\phi \in \Phi} (\theta_H I(\phi)) \leq V := \sup_{v \in \mathcal{V}} \int_{\Theta} F(\theta, v(\theta), \dot{v}(\theta)) d\theta, \quad (\mathcal{R})$$

where \mathcal{V} denotes the set of functions $v \in AC(\Theta; \mathbb{R})$ that satisfy Equation (17). Following the terminology of screening models, we refer to (\mathcal{R}) as the *relaxed problem* because it only takes into account the first-order condition induced by the incentive compatibility constraint (10). Our goal will be to show that at the optimum of the relaxed problem, this first-order condition is sufficient for incentive compatibility so that the solution to (\mathcal{P}) can be constructed from the solution to (\mathcal{R}) .

The relaxed problem (\mathcal{R}) is a scalar calculus of variations problem that can be solved using standard techniques (see, e.g., Mesterton-Gibbons (2009)). Specifically, using subscripts to denote partial derivatives, we have that a necessary condition for optimality is given by the Euler–Lagrange equation

$$F_{v(\theta)}(\theta, v(\theta), \dot{v}(\theta)) - \frac{d}{d\theta} F_{\dot{v}(\theta)}(\theta, v(\theta), \dot{v}(\theta)) = 0, \quad \theta \in \Theta, \quad (20)$$

and, because the boundary values of v are free, this second-order differential equation should be solved subject to the boundary conditions

$$F_{\dot{v}(\theta)}(\theta, v(\theta), \dot{v}(\theta)) = 0, \quad \theta \in \{0, \theta_H\}. \quad (21)$$

Calculating the derivatives involved in these expressions and simplifying the result leads to the boundary value problem

$$v(\theta)(1 + \ddot{v}(\theta)) - [\dot{v}(\theta)]^2 = \frac{3}{2\xi} (2v(\theta) - [\dot{v}(\theta)]^2)^{\frac{3}{2}}, \quad \theta \in \Theta \quad (22)$$

subject to

$$\dot{v}(0) = 0, \tag{23}$$

$$\theta_H = \xi \dot{v}(\theta_H) (2v(\theta_H) - [\dot{v}(\theta_H)]^2)^{-\frac{1}{2}}. \tag{24}$$

Our next result establishes the existence of a unique solution to this problem and verifies that it attains the supremum in the relaxed problem.

Proposition 3.4. *There exists a unique solution $v^* \in C^2(\Theta; \mathbb{R})$ to the boundary value problem defined by Equations (22)–(24). This solution is strictly increasing, strictly convex, and attains the supremum in the relaxed problem.*

3.3 | The optimal fund menu

By definition, we have that the value of the relaxed problem (\mathcal{R}) gives an upper bound on the value of the manager’s problem (\mathcal{P}). To show that these values coincide, and thus characterize the optimal menu, we need to construct a loading function $\phi^* \in \Phi$ that delivers to each investor the indirect utility prescribed for his type by the solution to the relaxed problem. The following theorem provides such a construction.

Theorem 3.5. *Denote by $v^* \in C^2(\Theta; \mathbb{R})$ the solution to the relaxed problem (\mathcal{R}) as defined in Proposition 3.4. Then,*

$$F^*(\theta) := F(\theta, v^*(\theta), \dot{v}^*(\theta)) > 0, \quad \theta \in \Theta, \tag{25}$$

and the loading function defined by

$$\phi^*(\theta) := \frac{1}{\alpha F^*(\theta)} \left(\sqrt{2v^*(\theta) - \dot{v}^*(\theta)^2}, \dot{v}^*(\theta) \right)^\top, \tag{26}$$

attains the supremum in (\mathcal{P}). In particular, every investor allocate a strictly positive amount in the fund targeted to his type.

Our last result in this section provides basic comparative statics for the optimal fund menu and the amount of fees that the manager receives from each type of investor.

Proposition 3.6. *The function $\phi_1^*(\theta)$ is decreasing in θ while the functions $F^*(\theta)$ and $\phi_2^*(\theta)$ are increasing.*

The proposition shows that the funds in the optimal menu become more tilted towards the non-familiar asset as θ increases and that the manager receives more fees from more optimistic investors. These results are intuitive. Indeed, investors with higher θ are more interested in the non-familiar asset. Knowing this, the manager gradually tilts the exposure of his funds towards the non-familiar risk to extract more fees from more optimistic investors. To guarantee that investors do not have any incentive to deviate from the fund targeted to them, the manager needs to

provide more optimistic investors with a disproportionately larger indirect utility than less optimistic investors. This is why the indirect utility function $v^*(\theta)$ induced by the optimal menu is not only strictly increasing, but also strictly convex in the investor's type.

Remark 3.7. Theorem 3.5 also characterizes the unique stationary optimal fund menu in a dynamic model with iid normal returns and CARA investors who do not update their beliefs. The assumption that investors do not learn seems somewhat stark in a dynamic setting but may be justified if the proportion of investors with the same beliefs remains constant over time even if the investors do change beliefs at the individual level, or if the manager keeps changing the features of the funds to roll back the learning of investors as in the obfuscation model of Carlin and Manso (2011).

4 | ANALYSIS

4.1 | Comparative statics

The comparative statics of equilibrium outcomes with respect to the risk premium ξ of the familiar asset and the range θ_H of perceived risk premia on the non-familiar asset are a lot more difficult to derive analytically. As shown by the following lemma, a notable exception concerns the manager's welfare.

Proposition 4.1. *The manager always prefers to face more optimistic investors in the sense that his value function M is increasing in both ξ and θ_H .*

The mechanism behind the above result is clear: as ξ or θ_H increase investors become more eager to invest in the risky assets and thus are willing to pay the manager a larger amount to get access to those assets through the funds.

The right panel of Figure 1 shows that the indirect utility of investors depends positively on the risk premium ξ of the familiar asset and negatively on the range θ_H of perceived risk premia on the non-familiar asset. The first result is partly due to the fact that an increase in ξ implies a reduction in the relative importance of the information friction and, thus, leads to a welfare increase. To understand the second result recall that, in order to satisfy the incentive compatibility constraint, the manager has to make sure that no investor has an incentive to under-report his type by switching to a fund that is targeted to less optimistic investors. Now fix an arbitrary type $\theta \in \Theta$. As the upper bound of the type space increases, the investors whose type are larger than θ and who have to be deterred from under-reporting their type as θ , become more optimistic on average. To prevent these investors from under-reporting the manager then needs to modify the menu to worsen the conditions he offers to investors of type θ and this explains why investors benefit from being part of a narrower customer base.

The left panel of Figure 1 illustrates the effect of a change in the range θ_H of perceived risk premia on the indirect utility of an investor who stands at a given percentile. As shown in the right panel, an increase in θ_H leads to a decrease in the utility of all investors. However, another effect comes into play when considering an investor at a given percentile because, as θ_H increases, the type and hence the indirect utility of investors at a given percentile also increase. As shown by the left panel of the figure, the first effect dominates at low percentiles while the second dominates at higher percentiles.

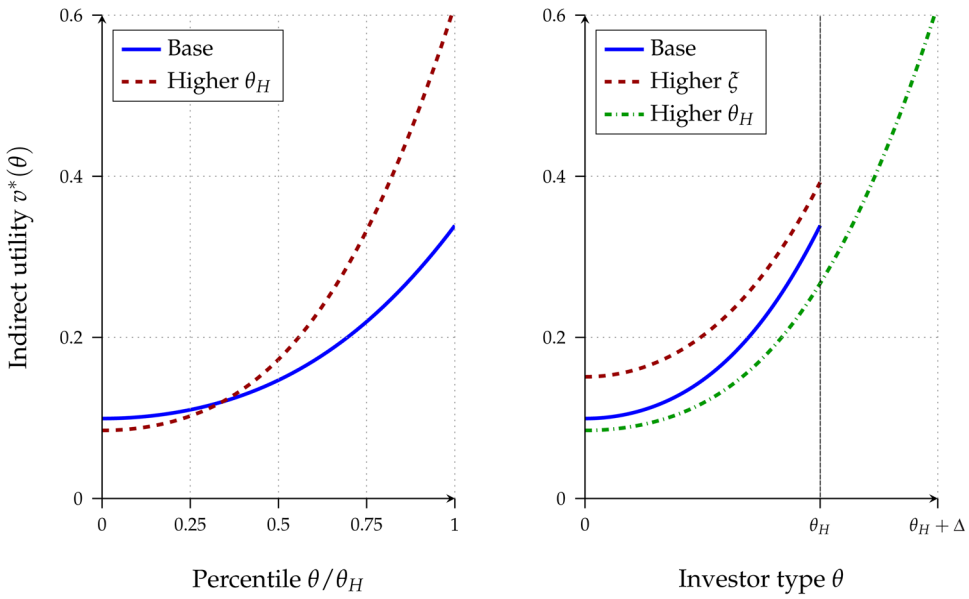


FIGURE 1 Effect of ξ and θ_H on the indirect utility of investors. [Color figure can be viewed at wileyonlinelibrary.com]

Notes: The solid curve plots the investor’s indirect utility as a function of his type (right) and his percentile in the distribution of types (left). The dashed curves illustrate how this utility changes when the range of perceived risk premia increases from θ_H to $\theta_H + \Delta$ while the dash-dotted curve indicates how it changes when the risk premium of the familiar asset increases

4.2 | Impact of the information friction

To elicit the impact of the information friction on the optimal menu and the welfare of the players, we now compare the outcomes of the model to those of the benchmark case where the manager knows the type of each individual investor but remains subject to a linear pricing constraint. The results in Appendix A.3 show that in this case, the optimal fund loading function $\phi^\circ(\theta)$, the optimal strategy $q^\circ(\theta)$ of investors of type θ , and their indirect utility $v^\circ(\theta)$ are given by

$$(\phi^\circ(\theta), q^\circ(\theta), v^\circ(\theta)) = \left(\frac{2\xi(\theta)}{\|\xi(\theta)\|^2}, \frac{\|\xi(\theta)\|^2}{4a}, \frac{\|\xi(\theta)\|^2}{8} \right). \tag{27}$$

Our first result compares the risk exposures offered by the manager to a given type of investor the two models.

Proposition 4.2. *The function*

$$\Delta(\theta) := \frac{\phi_2^\circ(\theta)}{\phi_1^\circ(\theta)} - \frac{\phi_2^*(\theta)}{\phi_1^*(\theta)} = \frac{\theta}{\xi} - \frac{v^*(\theta)}{g^*(\theta)}, \tag{28}$$

is nonnegative for all $\theta \in \Theta$ and such that $\Delta(0) = \Delta(\theta_H) = 0$.

The proposition shows that the lack of information regarding the beliefs of investors for the non-familiar asset leads the manager to offer funds that are more tilted towards the familiar asset

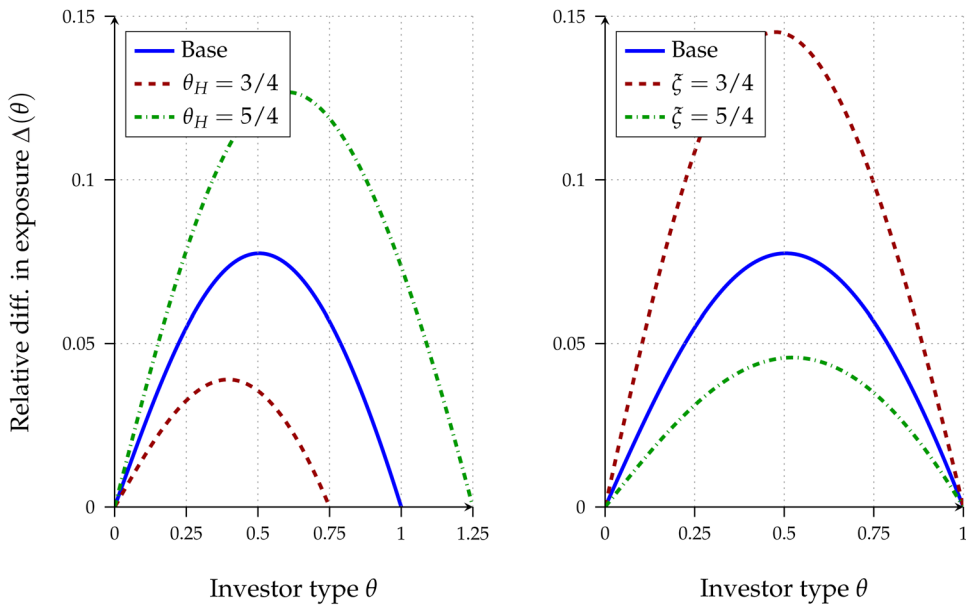


FIGURE 2 Relative difference in risk exposures compared to the first best. [Color figure can be viewed at wileyonlinelibrary.com]

Notes: The solid curves plots the function $\Delta(\theta)$ defined in Equation (28) at the base parameter values $\xi = \theta_H = 1$. The dashed curves illustrate the impact of a decrease in either θ_H (left) or ξ (right) while the dash-dotted curves illustrate the impact of an increase in these same parameters

than in the first best. This means that to achieve a given exposure in the non-familiar asset, a given investor needs to take a larger position in the familiar asset than he would have in the first best, and the manager uses the eagerness of investors to do so as a screening device. Our model thus provides a potential explanation for the fact that fund managers are often perceived as being home-biased because their funds are too geared towards familiar assets. Intuitively, the strength of this bias should be driven by the intensity of the information friction. Therefore, one expects that $\Delta(\theta)$ should increase as θ_H/ξ increases and Figure 2 confirms that this is the case.

The dome shape of the function $\Delta(\theta)$ that is apparent in both panels of Figure 2 is the result of two conflicting effects. On the one hand, the fact that more optimistic investors demand more of the non-familiar asset prompts the manager to intensify the distortions as θ increases to deter these more optimistic investors from underreporting their type. On the other hand, as θ increases, the mass of investors who might be tempted to underreport their type as being equal to θ becomes smaller and this implies that fewer distortions are needed to maintain incentive compatibility.

The second part of Proposition 4.2 shows that the risk exposures offered to the most pessimistic and most optimistic investors are the same as in the first best. However, this does not imply that these investors select the same allocation or receive the same utility as in the first best because, even though the risk exposures they are offered are the same, the prices that the manager demands for them are different.

Remark 4.3. The bias implied by Proposition 4.2 is partly due to the fact that all investors would invest the same amount in the familiar asset given direct access. Whether the result can be expected to hold in environments where both components of an investor's unconstrained demand vary with his type likely depends on the relative shapes of these demand components. For

example, if asset returns are correlated then the optimal unconstrained demand for asset 2 is increasing in type whereas the ratio of the demand for asset 1 to the demand for asset 2 is decreasing. As a result, the share of the non-familiar asset in the optimal unconstrained portfolio of an investor increases with his type as in our benchmark model and we thus expect the same bias to prevail.

Our next result provides a detailed comparison of the portfolio allocations and indirect utilities of investors in the two models.

Proposition 4.4. *There are types $\theta_1 \leq \bar{\theta} \leq \theta_2$ such that*

$$\begin{aligned} \{\theta \in \Theta : v^*(\theta) \leq v^\circ(\theta)\} &= [0, \bar{\theta}], \\ \{\theta \in \Theta : \pi(\theta, \phi^*(\theta))\phi_k^*(\theta) \leq q^\circ(\theta)\phi_k^\circ(\theta)\} &= [0, \theta_k], \quad k \in \{1, 2\}. \end{aligned} \quad (29)$$

The proposition shows that types below θ_1 invest less in both risky assets than in the first best case and suffer a utility loss, that types above θ_2 invest more in both risky assets than in the first best case and receive a utility gain, and that intermediate types in $[\theta_1, \theta_2]$ invest more in the familiar asset and less in the non-familiar asset than in the first best case. To understand these results, start by considering very low types. Since such investors lie at the bottom of the distribution, the manager needs to deter almost all other investors from pooling with them. To do so, he must offer them high prices and it naturally follows that such investors end up investing less in both assets and suffer a significant utility loss. More optimistic investors are offered better terms that lead them to invest more and to increase their exposure to the non-familiar asset. However, the discussion following Proposition 4.2 shows that their exposure to the familiar asset will increase at a faster rate, which explains why intermediate types invest more in the familiar asset and less in the non-familiar asset. Finally, as we approach the right tail of the distribution, the terms that are being offered to investors are so good that they invest more in both assets, and their indirect utility exceeds that of the first best.

When moving from the full information case to the asymmetric information case, the manager naturally suffers a decrease in utility since he now has less information. As θ_H increases, investors become more optimistic on average and thus more eager to trade the non-familiar asset. Therefore, the information friction becomes more intense and we thus expect the manager's utility loss to increase as a function of θ_H . Similarly, as the risk premium of the familiar asset increases the information friction becomes less intense because investors now tend to care more for the familiar asset and, as a result, we expect the manager's utility loss to decrease as a function of ξ . Figure 3 numerically confirms that these natural properties holds at the optimum of our model.

4.3 | Impact of the linear pricing constraint

To analyze the impact of the linear pricing constraint, we now compare the outcomes of our model to those of a model in which the manager is unconstrained in his choice of the pricing scheme. We show in Appendix A that in this case, the optimal pricing strategy is to offer a fixed cost equal to $\xi^2/(2a)$ for unlimited access to the familiar asset and a quantity-dependent unit price given by

$$\hat{p}(q) := \frac{1}{2}\theta_H - \frac{a}{4}q, \quad (30)$$

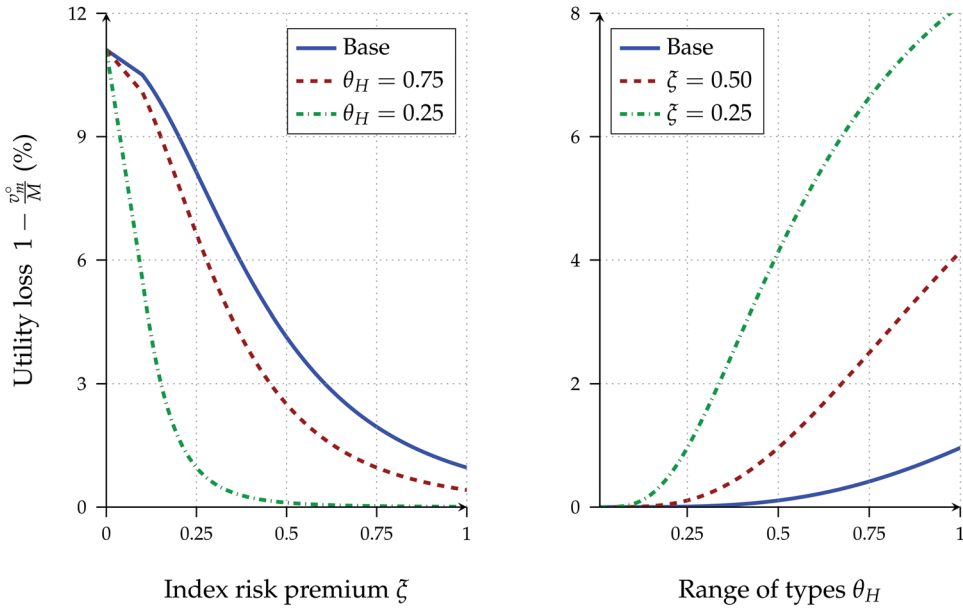


FIGURE 3 Manager’s utility loss compared to the first best. [Color figure can be viewed at wileyonlinelibrary.com]

Notes: The solid curves plot the relative utility loss of the manager as a function of the risk premium of the familiar asset and the range of perceived risk premia on the non-familiar asset. In each panel the dashed and dash-dotted curves illustrate the effect of a change in either the risk premium ξ of the familiar asset (right) or the range of investor types θ_H (left). The base case parameters used to construct this figure are $\xi = \theta_H = 1$

for trading the non-familiar asset. In response to this menu, an investor of type $\theta \in \Theta$ demands $\hat{q}_1(\theta) = \xi/a$ units of the familiar asset and $\hat{q}_2(\theta) = (2\theta - \theta_H)_+/a$ units of the non-familiar asset so that his expected utility is given by

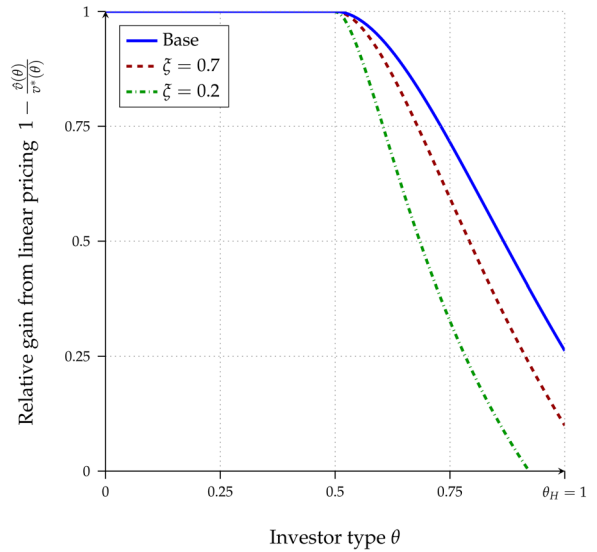
$$\hat{v}(\theta) := u\left(\theta, rw_0 + \hat{q}(\theta)^\top \varepsilon - \frac{\xi^2}{2a} - \hat{q}_2(\theta)\hat{p}(\hat{q}_2(\theta))\right) = \left(\theta - \frac{\theta_H}{2}\right)_+^2. \tag{31}$$

Comparing this solution to that of our model reveals two major differences.⁹ First, with nonlinear pricing investors of type $\theta \leq \theta_H/2$ who care less about the familiar asset than the average investor get zero utility which means that, in contrast to the linear pricing case, the manager is able to extract the whole surplus generated by investments in the familiar asset. This is intuitive. Indeed, because it is common knowledge that investors have identical preferences regarding the familiar asset, the manager knows exactly how many units each investor would want to acquire and thus can set his fixed price so as to extract the full surplus generated by this investment.

Second, and more importantly, nonlinear pricing makes it optimal for the manager to exclude investors who are less optimistic than average from the non-familiar asset market. By contrast, under linear pricing, the optimal menu is such that all investors hold the two risky assets and receive a strictly positive utility. This suggests that linear pricing improves the aggregate welfare of investors and may even result in individual gains for all investors if the benefit from recovering part of the surplus associated with the familiar asset is sufficient to offset the forgone quantity discounts implied by Equation (30) on the non-familiar asset. These intuitive properties seem

FIGURE 4 Utility gain from linear pricing. [Color figure can be viewed at wileyonlinelibrary.com]

Notes: This figure plots the relative difference between the indirect utility of investors in the main model and their indirect utility in the model where the manager can use nonlinear pricing. The solid curve represents gain from linear pricing in the base case where $\xi = \theta_H = 1$ while the dashed and dash-dotted curves illustrate the impact of a gradual decrease in the familiar asset risk premium



difficult to establish analytically. However, all our numerical simulations confirm that linear pricing indeed improves the aggregate welfare of investors, and Figure 4 illustrates that it may even lead to strict Pareto improvements when the ratio θ_H/ξ that measures the intensity of the information friction is low enough.

4.4 | The optimality of bundling

The normalization that we adopted in Proposition 3.1 implies that funds differ only in their exposure to the risky assets. This normalization is convenient for the derivation of the optimal menu but, in some cases, it may be more natural to instead normalize the funds in such a way that the optimal menu only includes all-equity funds that do not invest in the riskless asset. With this alternative normalization, the fund that targets investors of type $\theta \in \Theta$ is given by

$$(\gamma_{AE}(\theta), \phi_{AE}(\theta)) := \frac{[1, \phi^*(\theta)]}{\phi_1^*(\theta) + \phi_2^*(\theta)} = \frac{1}{v^*(\theta) + g^*(\theta)} \left[aF^*(\theta), \begin{pmatrix} g^*(\theta) \\ v^*(\theta) \end{pmatrix} \right], \tag{32}$$

where we have set

$$g^*(\theta) := \sqrt{2v^*(\theta) - [v^*(\theta)]^2}. \tag{33}$$

In particular, since $v^*(0) = 0$ by Theorem 3.5, we have $\phi_{AE}(0) = (1, 0)^\top$ so that the familiar asset is offered in the optimal menu with a fee rate given by

$$\gamma_1^* := \frac{aF^*(0)}{g^*(0)} = \xi - \sqrt{2v^*(0)} \in \left[\frac{1}{3}, \frac{2}{3} \right] \xi, \tag{34}$$

where the inclusion follows from the fact that, as we show in Appendix 3.3, the indirect utility of the most pessimistic investor lies in the interval $[\frac{1}{18}, \frac{2}{9}] \xi^2$.

This shows that the requirement of linear pricing leads the manager to engage in what Adams and Yellen (1976) refer to as *mixed* bundling – the familiar asset is available both separately and in packages. Note, however, that it is never optimal to also offer the non-familiar asset separately because this would prevent the manager from screening investors. Indeed, if the manager decides not to bundle, then the best he can do is to offer assets separately with fee rates

$$(\hat{\gamma}_1, \hat{\gamma}_2) := \operatorname{argmax}_{\gamma \in \mathbb{R}_+^2} \left\{ \gamma_1(\xi - \gamma_1)_+ + \int_0^{\theta_H} \gamma_2(\theta - \gamma_2)_+ \frac{d\theta}{\theta_H} \right\} = \left(\frac{\xi}{2}, \frac{\theta_H}{3} \right). \quad (35)$$

This pricing strategy in turn generates

$$M_0 := \hat{\gamma}_1(\xi - \hat{\gamma}_1)_+ + \frac{1}{\theta_H} \int_0^{\theta_H} \hat{\gamma}_2(\theta - \hat{\gamma}_2)_+ d\theta = \frac{1}{4}\xi^2 + \frac{2}{27}\theta_H^2, \quad (36)$$

in management fees and our next result confirms that this quantity is strictly lower than the amount of fees generated by the optimal fund menu.

Lemma 4.5. *It is never optimal to unbundle the assets, that is $M_0 < M$.*

4.5 | Exclusivity of funds

Consider now the *exclusive* case in which each investor can allocate capital to at most one fund. Given a menu satisfying properties 1 and 2 of Proposition 3.1, the optimal strategy of an investor of type $\theta \in \Theta$ who optimally allocates his wealth between fund $(1, \phi(\theta'))$ and the riskless asset is given by

$$\operatorname{argmax}_{q \in \mathbb{R}_+} u(\theta, rw_0 + q\mathcal{R}(1, \phi(\theta'))) = \pi(\theta, \phi(\theta')), \quad (37)$$

and delivers him the indirect utility

$$v(\theta, \theta') := u(\theta, rw_0 + \pi(\theta, \phi(\theta'))\mathcal{R}(1, \phi(\theta'))) = \frac{1}{2} \left(\frac{\phi(\theta')^\top \xi(\theta) - 1}{\|\phi(\theta')\|} \right)_+^2. \quad (38)$$

Under exclusivity, incentive compatibility only requires that each investor finds it optimal to pick the fund targeted to him in the sense that

$$v(\theta) = v(\theta, \theta) = \sup_{\theta' \in \Theta} v(\theta, \theta'), \quad \theta \in \Theta. \quad (39)$$

The Cauchy–Schwartz inequality guarantees that this condition is weaker than its non exclusive counterpart in Equation (10) so that the manager cannot do worse when the investors are forced to commit to a single fund. However, since

$$\left. \frac{dv(\theta, \theta')}{d\theta'} \right|_{\theta'=\theta} = \pi(\theta, \phi(\theta)) \left. \frac{d\alpha(\theta, \theta')}{d\theta'} \right|_{\theta'=\theta}, \quad (40)$$

we have that the first-order conditions induced by those two constraints coincide and it follows that the same menu is optimal under either constraint. As we will see below in Section 5.3, this result no longer holds when investors can directly access the familiar asset at a sufficiently low cost.

5 | DIRECT INVESTMENT IN THE FAMILIAR ASSET

In our main model, investors can only access the risky assets through the manager. We now relax this assumption by allowing them to directly access the familiar asset via an outside fund that charges an exogenous fee.¹⁰

5.1 | Formulation

Assume that investors can allocate capital to the riskless asset, the manager's funds, and an outside fund with net excess return $\mathcal{R}_1 := \varepsilon_1 - \gamma_1$ for some $\gamma_1 \in [0, \xi]$. In this case, the investors' budget constraint is

$$w_1(q, n, \mathbf{m}) := r\omega_0 + \int_{\mathcal{M}} \mathcal{R}(\gamma(m), \phi(m))q(dm) + n\mathcal{R}_1, \quad (41)$$

where $n \geq 0$ represents the amount invested in the outside fund. The optimization problem of an investor of type $\theta \in \Theta$ is then defined by

$$v_i(\theta, \mathbf{m}) := \sup_{(q,n) \in \mu_+(\mathcal{M}) \times \mathbb{R}_+} u(\theta, w_1(q, n, \mathbf{m})), \quad (42)$$

and the aggregation of the investors' decisions generates the amount of management fees given by Equation (6) where

$$(q^*(\cdot, \theta, \mathbf{m}), n^*(\theta, \mathbf{m})) = \operatorname{argmax}_{(q,n) \in \mu_+(\mathcal{M}) \times \mathbb{R}_+} u(\theta, w_1(q, n, \mathbf{m})), \quad (43)$$

denotes the best response of an investor of type $\theta \in \Theta$. To facilitate the analysis, we assume throughout that if the manager includes in his menu the familiar asset with a fee equal to γ_1 investors will direct their demand for the familiar asset to the manager rather than to the outside fund. Given this assumption, a menu is said to be optimal if it maximizes Equation (6) subject to Equation (43).

The following lemma elicits the conditions under which the presence of the outside fund has an impact on the optimal fund menu.

Lemma 5.1. *Assume that*

$$\gamma_1 > \gamma_1^* := \xi - \sqrt{2v^*(0)}, \quad (44)$$

where the function v^* is defined as in Proposition 3.4. Then the optimal fund menu is given by Theorem 3.5.

The intuition for the above result is clear: when confronted with two funds that offer the same risk exposure but different fee rates, investors will systematically discard the fund with the higher fee rate. Therefore, if γ_1 exceeds the fee rate γ_1^* that the manager offers as part of the optimal menu in the benchmark mode, then investors will stay away from the outside fund and the optimal fund menu will remain unchanged.

5.2 | The relaxed problem

Assume from now on that the condition of Lemma 5.1 fails. Proceeding as in Section 3 shows that the manager can without loss of generality focus on menus such that funds are indexed by types, fees are normalized to one, and each investor finds it optimal to only invest in the fund that target its type (see Appendix C.1). The incentive compatibility constraint is more stringent because the manager has to prevent investors from allocating capital not only to funds that do not target their type but also to the outside fund. The following result quantifies this observation by providing a characterization of incentive compatible loading functions.

Proposition 5.2. *A loading function $\phi : \Theta \rightarrow \mathbb{R}^2$ is incentive compatible given direct access to the familiar asset if and only if it satisfies condition (10) and*

$$\inf_{\theta \in \Theta} \left\{ \frac{\phi_1(\theta)}{\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1)_+ \right\} \geq \xi - \gamma_1, \tag{45}$$

in which case $\phi_1(\theta) > 0$ and $\phi(\theta)^\top \xi(\theta) > 1$ for all $\theta \in \Theta$.

Combining the above results shows that the optimal menu solves

$$M_1 := \sup_{\phi \in \Phi_1} I(\phi). \tag{P_1}$$

where $\Phi_1 \subseteq \Phi$ denotes the set of functions $\phi \in AC(\Theta; \mathbb{R}^2)$ that satisfy Equations (10) and (45). Following the logic of Lemma 3.3, our next result shows that the manager can use the indirect utility and marginal utility of investors as instruments.

Lemma 5.3. *Assume that $\phi \in \Phi_1$ is incentive compatible. Then, the indirect utility function defined by Equations (38) and (39) belongs to $AC(\Theta; \mathbb{R})$ and satisfies both Equation (18) and*

$$2v(\theta) \geq (\xi - \gamma_1)^2 + [\dot{v}(\theta)]^2, \tag{46}$$

for almost every type $\theta \in \Theta$.

The above lemma directly implies that

$$\theta_H M_1 \leq V_1 := \sup_{v \in \mathcal{V}_1} \int_{\Theta} F(\theta, v(\theta), \dot{v}(\theta)) d\theta, \tag{R_1}$$

where \mathcal{V}_1 denotes the set of functions $v \in AC(\Theta; \mathbb{R})$ that satisfy Equation (46). An important difference between this problem and (R) is that for the incentive compatibility condition (46) to

hold it is no longer sufficient that the integrand in the objective be real valued for all types. This implies that the usual sufficient optimality conditions need to be modified to explicitly account for the constraint and we show in Appendix C.2 how this can be achieved by constructing an appropriate Lagrangian.

To motivate the form of the solution consider investors with low types who do not care much for the non-familiar asset and have access to the familiar asset at a cheap price $\gamma_1 \leq \gamma_1^*$. To entice such investors to allocate capital to a fund that includes the non-familiar asset, the manager needs to offer them very favorable conditions that are likely to attract more optimistic investors and, thereby, fail the incentive compatibility constraint. We thus expect the constraint to bind over $[0, \theta^*]$ for some $\theta^* \in \Theta$ so that the manager offers only the familiar asset to sufficiently pessimistic investors. As the fee rate on the outside fund decreases, all investors become less willing to acquire exposure to the familiar asset otherwise than through the outside fund. This makes it gradually more difficult for the manager to screen investors by bundling and we expect that below a certain fee rate, it will no longer be optimal to do so. At that point the manager will pick a menu that is equivalent to offering the assets separately with fee rates γ_1 and $\hat{\gamma}_2 = \frac{1}{3}\theta_H$ (see Equation (35)). This menu delivers the indirect utility

$$s(\theta) := \sup_{q \in \mathbb{R}_+^2} u(\theta, rw_0 + q^\top(\varepsilon - \gamma)) = \frac{1}{2}(\xi - \gamma_1)^2 + \frac{1}{2}\left(\theta - \frac{\theta_H}{3}\right)_+^2, \quad (47)$$

and a direct calculation shows that this function satisfies condition (46) with an equality. Based on this observation, we conjecture that when γ_1 is sufficiently low, the constraint binds not only for low types but throughout the type space.

The following proposition confirms the above conjectures and provides a complete solution to the relaxed problem.

Proposition 5.4.

- (a) If $\gamma_1 \leq \frac{1}{3}\xi$, then $v_1^*(\theta) := s(\theta)$ attains the supremum in (\mathcal{R}_1) .
 (b) Assume $\gamma_1 \in (\frac{1}{3}\xi, \gamma_1^*]$ and denote by $(w, \theta^*) \in C_p^2(\Theta; \mathbb{R}) \times \Theta$ the unique solution to the free boundary problem defined by

$$w(\theta)(1 + \ddot{w}(\theta)) = [\dot{w}(\theta)]^2 + \frac{3}{2\xi}(2w(\theta) - [\dot{w}(\theta)]^2)^{\frac{3}{2}}, \quad (48)$$

subject to

$$0 = \dot{w}(\theta^*) = w(\theta^*) - \frac{1}{2}(\xi - \gamma_1)^2 \quad (49)$$

$$= \theta_H - \xi \dot{w}(\theta_H)(2w(\theta_H) - [\dot{w}(\theta_H)]^2)^{-\frac{1}{2}}. \quad (50)$$

Then, the function

$$v_1^*(\theta) := \mathbf{1}_{\{\theta \leq \theta^*\}} \frac{1}{2}(\xi - \gamma_1)^2 + \mathbf{1}_{\{\theta > \theta^*\}} w(\theta), \quad (51)$$

attains the supremum in (\mathcal{R}_1) .

5.3 | The optimal fund menu

As in our main model, the value of the relaxed problem (\mathcal{R}_1) gives an upper bound on the value of the actual problem (\mathcal{P}_1) and, to show that these values coincide, we need to construct an incentive compatible loading function that delivers to each investor the indirect utility prescribed for his type by the solution to the relaxed problem. This is the content of the following:

Theorem 5.5. *Assume $\gamma_1 \leq \gamma_1^*$ and denote by $v_1^* \in C_p^2(\Theta; \mathbb{R})$ the solution to the relaxed problem (\mathcal{R}_1) as defined in Proposition 5.4. Then*

$$\inf_{\theta \in \Theta} F(\theta, v_1^*(\theta), \dot{v}_1^*(\theta)) > 0, \quad (52)$$

and the fund loading function

$$\phi_1^*(\theta) := \frac{1}{\alpha F(\theta, v_1^*(\theta), \dot{v}_1^*(\theta))} \left(\sqrt{2v_1^*(\theta) - [v_1^*(\theta)]^2}, v_1^*(\theta) \right)^\top, \quad (53)$$

attains the supremum in (\mathcal{P}_1). In particular, $\gamma_1 \phi_1^*(\theta) = (1, 0)^\top$ for all $\theta \leq \theta^*$, so that the manager only offers the familiar asset to all sufficiently low types.

A comparison of Theorems 3.5 and 5.5 reveals two important differences. First, the presence of competition from the outside fund forces the manager to offer the familiar asset at the market rate γ_1 rather than at the monopolistic fee rate γ_1^* of Lemma 5.1. In addition, the optimal menu is such that the manager offers the familiar asset not only to the most pessimistic investors but to a group of sufficiently pessimistic investors who, therefore, find themselves excluded from the non-familiar asset market.

Second, if the competition induced by the outside fund is sufficiently fierce, then it is no longer useful for the manager to bundle assets and the optimum can be implemented by offering the familiar asset at the market rate γ_1 and the non-familiar asset at rate $\frac{1}{3}\theta_H$. Note, however, that this conclusion is fragile and dependent on our specific assumptions. In our model, the introduction of a cheap outside fund effectively reduces the optimal number of funds from a continuum to only two but this only occurs because there are two risky assets. In a general model with $n \geq 3$ assets, we conjecture that unbundling is unlikely to occur unless the manager holds a monopoly on a single risky asset and faces fierce competition on all others.

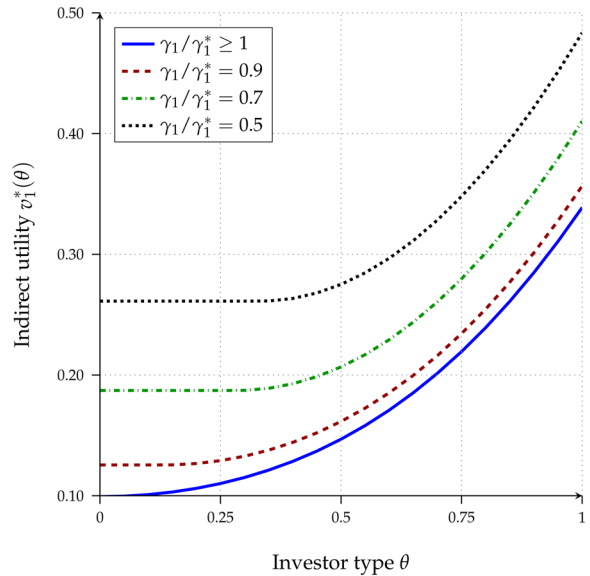
Since the fee rate γ_1 on the familiar asset is lower than the monopolistic rate γ_1^* , it follows from Equation (34) that we have $v^*(0) \leq v_1^*(0)$. This shows that the presence of an outside fund improves the welfare of the most pessimistic investors who are those that care the most for the familiar asset. As illustrated by Figure 5, this property actually holds for all investors because the presence of competition combined with the need to maintain incentive compatibility forces the manager to offer better terms not only to the most pessimistic investors but to all of them.

To conclude, let us briefly examine the non-exclusivity of funds in the presence of an outside fund. If the manager has the ability to commit each investor to a single fund, then incentive compatibility and individual rationality require that the investors' indirect utility satisfies Equation (39) and the participation constraint

$$\inf_{\theta \in \Theta} v(\theta) \geq \frac{1}{2}(\xi - \gamma_1)^2, \quad (54)$$

FIGURE 5 Indirect utility in the presence of an outside fund. [Color figure can be viewed at wileyonlinelibrary.com]

Notes: This figure plots the indirect utility of an investor as a function of his type for different levels of the market fee rate γ_1 on the outside fund



which states that optimally picking one fund out of the menu is preferable to optimally investing in the outside fund. Our last result provides a complete solution to the model in this case and shows that the ability to commit investors to a single fund may have value when some measure of competition is introduced in the model.

Proposition 5.6. *Assume that the manager can commit each investor to a single fund. Then, the indirect utility of investors and the optimal loading function are given by Equations (51) and (53) for all $\gamma_1 \leq \gamma_1^*$. As a result, the ability to commit investors has strictly positive value to the manager if and only if $\gamma_1 < \xi/3$.*

6 | CONCLUSION

In this paper, we argue that offering a menu of funds is optimal for an investment firm that has incomplete information about the characteristics of its customer base and is required to use linear pricing. To illustrate this mechanism, we study the optimal offering strategy of a manager who is constrained to use fraction-of-fund fees and does not observe the beliefs of investors regarding one of the risky assets. We show that the optimal menu can be explicitly constructed from the solution to a calculus of variations problem that optimizes over the indirect utility that investors receive. We provide a complete characterization of the optimal menu and study its most salient features. In particular, we show that the information friction leads the manager to offer funds that are inefficiently tilted towards the asset that is not subject to the information friction, and argue that this result provides a novel information-based explanation for the home bias.

While the tractability of our model rests on specific assumptions regarding the beliefs, preferences, and heterogeneity of investors, we believe that some of our key conclusions are not dependent on these assumptions. Instead of (or in addition to) differing in their beliefs, the investors could also differ along other dimensions such as risk aversion, initial endowments, hedging needs, the assets they are willing to invest in, or the risks they are exposed to prior to choosing their portfolio allocation. Such generalizations are likely to significantly complicate the analysis of the

model, but we conjecture that they would not undermine our qualitative message. In particular, we expect that in these more general environments, the linear pricing constraint will still induce the manager to bundle assets into funds as a way of screening investors. We leave these challenging extensions of our basic framework for future research.

ACKNOWLEDGMENTS

This work was initiated during a visit of Jakša Cvitanić to EPFL in the context of the special semester on *Stochastic Dynamic Models in Mathematical Finance, Econometrics, and Actuarial Sciences* organized by the Centre International Bernoulli in the Spring of 2017. Financial support from the Swiss Finance Institute (Hugonnier) and from the NSF under grant DMS-1810807 (Cvitanić) is gratefully acknowledged. We thank Jonathan Berk, Ryan Donnelly, Federico Echenique, Ruediger Fahlenbrach, Damir Filipović, Benjamin Hebert, Daniel Kuhn, Semyon Malamud, Yaroslav Melnyk, Erwan Morellec, Chris Rogers, and seminar participants at Cambridge, the University of Southern California, Boston University, and Caltech for conversations on the topic of this paper.

ENDNOTES

- ¹ Morningstar (2018) reports that the 150 largest US fund families jointly offer 7687 funds of which 4706 are equity funds, and that the average number of funds offered by a given family in that group is equal to 51 with some sponsors, such as Fidelity, offering more than 400 different funds.
- ² The 1970 Amendment to the Investment Advisors Act of 1940 allows managers of mutual funds listed in the US to use performance fees only if they are symmetric around a benchmark. As a result, the vast majority of funds in US use linear schedules known as fraction-of-fund fees. For example, Das and Sundaram (1998) report that as of 1998, only 1.4% of funds used performance fees.
- ³ See for example Mussa and Rosen (1978), and Wilson (1993) or Laffont and Martimort (2009) for a textbook treatment. An explicit derivation of the optimal nonlinear pricing scheme in the setting of our mutual fund model is provided in Appendix A.2.
- ⁴ See Cooper (2013) for a survey of the literature on the home bias and Hau and Rey (2008) for a study of the home bias at the mutual fund level in which the authors show that, while of lower magnitude than among other investor classes, the home bias is nonetheless present in the decision of equity mutual fund managers.
- ⁵ The multiplication by the risk aversion $\alpha > 0$ and the subtraction of the constant term arw_0 in the definition of the investors' preferences is without loss of generality and allows to simplify many expressions throughout the text and appendix.
- ⁶ The fact that investors have to go through the manager to access the risky assets may be due to transaction and/or informational costs. For example, one may think that investors form their beliefs about the non-familiar asset using information provided as part of the onboarding process and that this information would not be otherwise available to investors.
- ⁷ In practice, management fees are calculated on the basis of the net asset value at the end rather than at the beginning of the period and thus include a form of performance sensitivity. We focus on the case where fees computed at the beginning of the period to avoid the nonlinearity that may arise from the fact that with possibly unbounded excess returns the end of period asset value may be negative.
- ⁸ See for example Mussa and Rosen (1978) and Laffont and Martimort (2009) for a textbook treatment. A derivation of the optimal nonlinear pricing scheme in the setting of our delegated portfolio management model is provided in Appendix A.2.
- ⁹ Note that this pricing scheme is not unique. Instead of unbundling the assets, the manager could induce the same amount of fees and the same investor utilities by offering a menu of *mutually exclusive* funds with loadings $\phi(\theta) = \hat{q}(\theta)$ and unit price $\xi^2/(2a) + \hat{q}_2(\theta)\hat{p}(\hat{q}_2(\theta))$ to which each investor can allocate either zero or one unit of capital.
- ¹⁰ We focus on the case where the outside fund coincides with the familiar asset risk as it is the most natural in our context. However, a qualitatively similar solution applies if, instead of the familiar asset, investors can directly access the unfamiliar asset.

REFERENCES

- Adams, W., & Yellen, J. (1976). Commodity bundling and the burden of monopoly. *Quarterly Journal of Economics*, 90(3), 475–498.
- Admati, A., & Pfleiderer, P. (1997). Does it all add up? Benchmarks and the compensation of active portfolio managers. *Journal of Business*, 70(3), 323–350.
- Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica*, 64(1), 51–75.
- Basak, S., & Pavlova, A. (2013). Asset prices and institutional investors. *American Economic Review*, 103(5), 1728–58.
- Basak, S., Pavlova, A., & Shapiro, A. (2007). Optimal asset allocation and risk shifting in money management. *Review of Financial Studies*, 20, 1583–1621.
- Basov, S. (2006). *Multidimensional screening*. Studies in economic theory. Springer. Retrieved from <https://books.google.ch/books?id=wX0AUHFusecC>.
- Berk, J., & Green, R. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112(6), 1269–1295.
- Berk, J., Van Binsbergen, J., & Liu, B. (2017). Matching capital and labor. *Journal of Finance*, 72(6), 2467–2504.
- Bhattacharya, S., & Pfleiderer, P. (1985). Delegated portfolio management. *Journal of Economic Theory*, 36(1), 1–25.
- Breton, M., Hugonnier, J., & Masmoudi, T. (2010). Mutual fund competition in the presence of dynamic flows. *Automatica*, 46, 1176–1185.
- Brown, D., & Wu, Y. (2016). Mutual fund flows and cross-fund learning within families. *Journal of Finance*, 71(1), 383–424.
- Carlin, B., & Manso, G. (2011). Obfuscation, learning, and the evolution of investor sophistication. *Review of Financial Studies*, 24, 754–785.
- Carpenter, J. (2000). Does option compensation increase managerial risk appetite? *Journal of Finance*, 55(5), 2311–2331.
- Cooper, I. (2013). The equity home bias puzzle: A survey. *Foundations and Trends in Finance*, 7, 289–416. <https://doi.org/10.1561/05000000039>
- Cuoco, D., & Kaniel, R. (2011). Equilibrium prices in the presence of delegated portfolio management. *Journal of Financial Economics*, 101(2), 264–296.
- Cvitanić, J., & Xing, H. (2018). Asset pricing under optimal contracts. *Journal of Economic Theory*, 173, 142–180.
- Dangl, T., Wu, Y., & Zechner, J. (2008). Market discipline and internal governance in the mutual fund industry. *Review of Financial Studies*, 21(5), 2307–2343.
- Das, S., & Sundaram, R. (1998). *On the regulation of fee structures in mutual funds* [Working paper]. National Bureau of Economic Research.
- Das, S., & Sundaram, R. (2002). Fee speech: Signaling, risk-sharing, and the impact of fee structures on investor welfare. *Review of Financial Studies*, 15(5), 1465–1497.
- Dybvig, P., Farnsworth, H., & Carpenter, J. (2010). Portfolio performance and agency. *Review of Financial Studies*, 23(1), 1–23.
- Gruber, M. (1996). Another puzzle: The growth in actively managed mutual funds. *Journal of Finance*, 51(3), 783–810.
- Hau, H., & Rey, H. (2008). Home bias at the fund level. *The American Economic Review*, 98(2), 333–338.
- Hirsch, M., Smale, S., & Devaney, R. (2013). *Differential equations, dynamical systems, and an introduction to chaos* (3rd ed.). Elsevier/Academic Press.
- Hugonnier, J., & Kaniel, R. (2010). Mutual fund portfolio choice in the presence of dynamic flows. *Mathematical Finance*, 20(2), 187–227.
- Khorana, A., & Servaes, H. (2012). What drives market share in the mutual fund industry? *Review of Finance*, 16, 81–113.
- Laffont, J.-J., & Martimort, D. (2009). *The theory of incentives: The principal-agent model*. Princeton University Press.
- Luenberger, D. (1969). *Optimization by vector space methods*. Wiley.
- Lynch, A., & Musto, D. (2003). How investors interpret past fund returns. *Journal of Finance*, 58(5), 2033–2058.
- Mamaysky, H., & Spiegel, M. (2002). *A theory of mutual funds: Optimal fund objectives and industry organization* [Working paper].
- Massa, M. (2000). *Why so many mutual funds? Mutual fund families, market segmentation and financial performance* [INSEAD Working paper].

- McAfee, P., McMillan, J., & Whinston, M. (1989). Multiproduct monopoly, commodity bundling, and correlation of values. *Quarterly Journal of Economics*, 104(2), 371–383.
- Mesterton-Gibbons, M. (2009). *A primer on the calculus of variations and optimal control theory*, volume 50 of *Student Mathematical Library*. American Mathematical Society.
- Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38, 175–208.
- Morningstar. (2018). *Fund family 150: Due diligence on the 150 largest fund families*. Morningstar due diligence reports.
- Müller, M. (1927). Über das fundamentaltheorem in der theorie der gewöhnlichen differentialgleichungen. *Mathematische Zeitschrift*, 26(1), 619–645.
- Mussa, M., & Rosen, S. (1978). Monopoly and product quality. *Journal of Economic Theory*, 18(2), 301–317.
- Ou-Yang, H. (2003). Optimal contracts in a continuous-time delegated portfolio management problem. *Review of Financial Studies*, 16(1), 173–208.
- Pastor, L., & Stambaugh, R. (2012). On the size of the active management industry. *Journal of Political Economy*, 120(4), 740–781.
- Rochet, J.-C. (1985). The taxation principle and multi-time Hamilton–Jacobi equations. *Journal of Mathematical Economics*, 14(2), 113–128.
- Rochet, J.-C. (1987). A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics*, 16(2), 191–200.
- Rochet, J.-C., & Choné, P. (1998). Ironing, sweeping, and multidimensional screening. *Econometrica*, 66(4), 783–826.
- Rothschild, C. (2015). Nonexclusivity, linear pricing, and annuity market screening. *Journal of Risk and Insurance*, 82(1), 1–32.
- Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2014). *Lectures on stochastic programming* (2nd ed.). Society for Industrial and Applied Mathematics.
- Spence, M. (1980). Multi-product quantity-dependent prices and profitability constraints. *Review of Economic Studies*, 47(5), 821–841.
- Stole, L. (2001). *Lectures on the theory of contracts and organizations*. [Lectures notes], University of Chicago.
- Wilson, R. (1993). *Nonlinear pricing*. Oxford University Press.

How to cite this article: Cvitanić J, Hugonnier J. Optimal fund menus. *Mathematical Finance*. 2022;32:455–516. <https://doi.org/10.1111/mafi.12341>

APPENDIX A: BENCHMARK CASES

A.1 | Free access to all assets

If investors can freely trade all existing assets then the indirect utility of the manager is zero, and an investor of type $\theta \in \Theta$ chooses his optimal investment by solving

$$\bar{v}(\theta) := \sup_{q \in \mathbb{R}^2} u(\theta, rw_0 + \mathcal{R}(0, q)) = \sup_{q \in \mathbb{R}^2} \left\{ aq^\top \xi(\theta) - \frac{a^2}{2} \|q\|^2 \right\}. \quad (\text{A.1})$$

The unique solution to this concave problem is $\bar{q}(\theta) := \frac{1}{a} \xi(\theta)$. Substituting this solution back into the objective shows that the investor's indirect utility is

$$\bar{v}(\theta) = a\bar{q}(\theta)^\top \xi(\theta) - \frac{a^2}{2} \|\bar{q}(\theta)\|^2 = \frac{1}{2} \|\xi(\theta)\|^2. \quad (\text{A.2})$$

This function naturally constitutes an upper bound on the indirect utility that investors can hope to achieve when they no longer have direct access to all assets.

A.2 | Nonlinear pricing

Assume now that the manager does not observe the types of investors, but is allowed to use any price schedule that is a function of the amount invested. In this case, an investor of type $\theta \in \Theta$ chooses his allocation by solving

$$\hat{v}(\theta) := \sup_{q \in \mathbb{R}_+^2} u(\theta, rw_0 + \mathcal{R}(0, q) - P(q)), \quad (\text{A.3})$$

where $P(q)$ is the fee that the manager charges for $q \in \mathbb{R}_+^2$ and, denoting the solution by $\hat{q}(\theta)$, we have that the manager's problem can be formulated as

$$\hat{M} := \sup_P \int_{\Theta} P(\hat{q}(\theta)) \frac{d\theta}{\theta_H}. \quad (\text{A.4})$$

Since

$$P(\hat{q}(\theta)) = \hat{q}(\theta)^\top \xi(\theta) - \frac{a}{2} \|\hat{q}(\theta)\|^2 - \frac{\hat{v}(\theta)}{a}, \quad (\text{A.5})$$

we have that the manager's objective function can be written as

$$\begin{aligned} \int_{\Theta} P(\hat{q}(\theta)) \frac{d\theta}{\theta_H} &= \int_{\Theta} \left(\hat{q}(\theta)^\top \xi(\theta) - \frac{a}{2} \|\hat{q}(\theta)\|^2 - \frac{\hat{v}(\theta)}{a} \right) \frac{d\theta}{\theta_H} \\ &= \int_{\Theta} \left(\hat{q}(\theta)^\top \xi(\theta) - \frac{a}{2} \|\hat{q}(\theta)\|^2 - \frac{1}{a} \frac{d\hat{v}(\theta)}{d\theta} (\theta_H - \theta) \right) \frac{d\theta}{\theta_H} - \frac{\hat{v}(0)}{a} \\ &= \int_{\Theta} \left(\hat{q}_1(\theta) \xi + \hat{q}_2(\theta) (2\theta - \theta_H) - \frac{a}{2} \|\hat{q}(\theta)\|^2 - \frac{\hat{v}(0)}{a} \right) \frac{d\theta}{\theta_H}, \end{aligned} \quad (\text{A.6})$$

where the second equality follows from integration by parts, and the third follows from the envelope condition, which requires that

$$\frac{d\hat{v}}{d\theta}(\theta) = \frac{d}{d\theta} \left\{ a(q^\top \xi(\theta) - P(q)) - \frac{a^2}{2} \|q\|^2 \right\}_{q=\hat{q}(\theta)} = a\hat{q}_2(\theta). \quad (\text{A.7})$$

Maximizing under the integral sign shows that whenever possible the manager should choose the price function in such a way that $\hat{v}(0) = 0$ and

$$\hat{q}(\theta) = \frac{\mathbf{e}_1}{a} \xi + \frac{\mathbf{e}_2}{a} (2\theta - \theta_H)_+, \quad (\text{A.8})$$

where $(\mathbf{e}_1, \mathbf{e}_2)$ denotes the orthonormal basis of \mathbb{R}^2 . Since the above allocation is weakly increasing, it follows from the taxation principle (see for example Laffont and Martimort (2009)) that

there exists a price function $\hat{P}(q)$ that implements it, in the sense that

$$\hat{q}(\theta) \in \operatorname{argmax}_{q \in \mathbb{R}_+^2} \left\{ a(q^\top \xi(\theta) - \hat{P}(q)) - \frac{a^2}{2} \|q\|^2 \right\}. \quad (\text{A.9})$$

The indirect utility of an investor of type $\theta \in \Theta$ can be computed from the envelope condition which requires that

$$\hat{v}(\theta) = \hat{v}(0) + \int_0^\theta d\hat{v}(x) = \int_0^\theta a\hat{q}_2(x)dx = \left(\theta - \frac{1}{2}\theta_H \right)_+^2. \quad (\text{A.10})$$

Using this formula in conjunction with Equation (A.5) then shows that the amount of fees that the managers receives from an investor of type $\theta \in \Theta$ is

$$\hat{P}(\hat{q}(\theta)) = \frac{1}{2a}\xi^2 + \frac{1}{a}(2\theta - \theta_H)_+ \left(\frac{3}{4}\theta_H - \frac{1}{2}\theta \right), \quad (\text{A.11})$$

and it now follows from the taxation principle (see Rochet (1985)) that a price function, which allows the manager to achieve his optimum is given by

$$\hat{P}(q) = \mathbf{1}_{\{q_1 \neq 0\}} \frac{1}{2a}\xi^2 + q_2 \left(\frac{1}{2}\theta_H - \frac{a}{4}q_2 \right). \quad (\text{A.12})$$

The interpretation of the above results is clear. All investors are able to achieve their optimal level of exposure to the familiar asset (i.e., $\hat{q}_1(\theta) = \tilde{q}_1(\theta)$), but the fact that the manager is perfectly informed about investors' beliefs regarding this asset allows him to fully extract the corresponding surplus. On the other hand, because the manager does not observe the investors' beliefs regarding the non-familiar asset, he must screen them along this dimension. To this end, he uses a price schedule that is strictly concave so that marginal prices decrease with quantities. In equilibrium, all investors except those of the highest type achieve inefficiently small exposures to the non-familiar asset, and investors who are less optimistic than average even get excluded from that market because serving such investors would require lower prices that would in turn reduce the fees that the manager extracts from more optimistic investors.

Importantly, the nonlinear pricing solution cannot be implemented through a menu of nonexclusive linearly priced funds with different fees and risk exposures. If it could then this implementation would also provide an optimal menu for our main model and, as a result, the linear pricing constraint would have no effect.

Remark A.1. The simple form of the optimal nonlinear pricing scheme in Equation (A.12) is due to the assumption of independent returns, which effectively results in a form of separation whereby the manager extracts the whole surplus on the familiar asset and uses a standard nonlinear pricing scheme (see e.g., Basov (2006, Example 173, p.127)) for the asset that is subject to the information friction. If the independence assumption is removed then the demand of an investor for both assets will be a function of his type and, as a result, the optimal pricing scheme will take a much more complicated form.

Remark A.2. Approximating the quantity dependent term $(\theta_H/2 - aq_2/4)$ by a piecewise constant function shows that the nonlinear pricing solution can be approached through a menu of mutually exclusive linearly priced bundles that are subject to minimum and maximum investment requirements. In this sense the nonlinear pricing scheme – which is best interpreted as a broker – can still resemble a fund family if the underlying assets are viewed as funds.

A.3 | Linear pricing under complete information

Assume now that the manager is fully informed about the type of each investor and interacts with each of them in a bilateral way, but is still required to use linear price schedules. If the manager offers a fund (γ, ϕ) to such an investor of type $\theta \in \Theta$ then the amount that this investor will optimally allocate to the fund is

$$\begin{aligned} \operatorname{argmax}_{q \in \mathbb{R}_+} u(\theta, rw_0 + q\mathcal{R}(\gamma, \phi)) &= \operatorname{argmax}_{q \in \mathbb{R}_+} \left\{ aq(\phi^\top \xi(\theta) - \gamma) - \frac{a^2}{2} \|q\phi\|^2 \right\} \\ &= \frac{1}{a\|\phi\|^2} (\phi^\top \xi(\theta) - \gamma)_+, \end{aligned} \quad (\text{A.13})$$

where the second equality follows from an application of the Kuhn–Tucker conditions. Taking this best response into account the manager then solves

$$v_m^\circ(\theta) := \sup_{(\gamma, \phi) \in \mathbb{R} \times \mathbb{R}^2} \frac{\gamma}{a\|\phi\|^2} (\phi^\top \xi(\theta) - \gamma)_+. \quad (\text{A.14})$$

Since the objective function only depends on the vector $\nu = \phi/\gamma$, we may without loss of generality normalize the fee rate to 1. With this normalization,

$$v_m^\circ(\theta) = \sup_{\nu \in \mathbb{R}^2} \frac{1}{a\|\nu\|^2} (\nu^\top \xi(\theta) - 1)_+, \quad (\text{A.15})$$

and solving that problem shows that the linearly priced fund that the manager offers to investors of type $\theta \in \Theta$ is given by $(1, \phi^\circ(\theta))$ with

$$\phi^\circ(\theta) := \operatorname{argmax}_{\nu \in \mathbb{R}^2} \frac{1}{a\|\nu\|^2} (\nu^\top \xi(\theta) - 1)_+ = \frac{2\xi(\theta)}{\|\xi(\theta)\|^2}. \quad (\text{A.16})$$

Substituting into Equations (A.13) and (A.15) shows that the amount that investors of type $\theta \in \Theta$ allocate to the fund, their indirect utility, and the manager's indirect utility are explicitly given by

$$(q^\circ(\theta), v^\circ(\theta)) := \left(\frac{\|\xi(\theta)\|^2}{4a}, \frac{\|\xi(\theta)\|^2}{8} \right), \quad (\text{A.17})$$

and

$$v_m^\circ := \int_{\Theta} v_m^\circ(\theta) \frac{d\theta}{\theta_H} = \frac{1}{4a} \left(\xi^2 + \frac{1}{3} \theta_H^2 \right). \quad (\text{A.18})$$

The most salient features of this solution can be summarized as follows. First, and contrary to what happens when all price schedules are allowed, the indirect utility of investors depends on the risk premium ξ of the familiar asset. This shows that linear pricing prevents the manager from

extracting the whole surplus associated with the familiar asset. Second, each investor allocates a strictly positive amount to the fund that the manager proposes to him. Third, the optimal exposure of an investor of type $\theta \in \Theta$ to the two risky assets are given by

$$q_2^\circ(\theta)\phi_2^\circ(\theta) = \frac{\xi(\theta)}{2a}. \tag{A.19}$$

Because investors have quadratic preferences and agree on the familiar asset risk premium, their optimal exposures to the familiar asset are the same, but more optimistic investors naturally choose a larger exposure to the non-familiar asset. The relative composition of the investors' optimal portfolios

$$\frac{q^\circ(\theta)\phi_2^\circ(\theta)}{q^\circ(\theta)\phi_1^\circ(\theta)} = \frac{\tilde{q}_2(\theta)}{\tilde{q}_1(\theta)} = \frac{\theta}{\xi}, \tag{A.20}$$

is the same as in the case where investors can freely trade the risky assets, but the overall risk exposure

$$\|q^\circ(\theta)\phi^\circ(\theta)\| = \frac{1}{2a}\|\xi(\theta)\| = \frac{1}{2}\|q_0(\theta)\|, \tag{A.21}$$

is smaller by a factor of two. The intuition for this result is that, since the loadings vector is determined only up to a multiplicative constant, offering the optimal fund $(1, \phi^\circ(\theta))$ to an investor of type θ is equivalent to offering him the fund $\frac{1}{a}\xi(\theta)$ that he would have picked on his own but with a fee equal $\frac{1}{2a}\|\xi(\theta)\|^2$ and, given this fee, the investor's optimal strategy is to invest half of what he would have on his own. Fourth, and last, the amount that an investor allocates to the fund (and hence the amount of fees that he pays) is increasing in both his risk tolerance and his type. As a result, the manager collects more fees from more optimistic or less risk averse investors, and his utility is increasing in $1/a$, ξ , and the parameter θ_H that determines the average belief of investors.

APPENDIX B: PROOF OF THE RESULTS IN SECTION 3

B.1 | The revelation principle

Before proceeding with the proof of Proposition 3.1 we start by establishing some useful results about the investors' problem (5).

Lemma B.1. *The measure $q^* \in \mu_+(\mathcal{M})$ is optimal for an investor of type $\theta \in \Theta$ if and only if it satisfies*

$$\int_{\mathcal{M}} \left\{ \phi(m)^\top \left(\xi(\theta) - a \int_{\mathcal{M}} \phi(n)q^*(dn) \right) - \gamma(m) \right\} \nu(dm) \leq 0, \tag{B.1}$$

for all measures ν in the set

$$\mathcal{F}(q^*) := \{ \nu \in \mu(\mathcal{M}) : \exists \delta > 0 \text{ such that } q^* + \delta\nu \in \mu_+(\mathcal{M}) \}. \tag{B.2}$$

In particular, the null measure is optimal for investors of type $\theta \in \Theta$ if and only if the menu is such that $\phi(m)^\top \xi(\theta) \leq \gamma(m)$ for all $m \in \mathcal{M}$.

Proof. The first part follows from standard results on convex optimization in infinite dimensional spaces, see for example Luenberger (1969, Chapter 7). The second part follows from the first part by taking $q^* = 0$ and observing that $\mathcal{F}(0) = \mu_+(\mathcal{M})$. \square

Lemma B.2. *The value function of an investor of type $\theta \in \Theta$ satisfies*

$$v_i(\theta, \mathbf{m}) = \sup_{q \in \mu_{2,+}(\mathcal{M})} u(\theta, w_1(q, \mathbf{m})) \tag{B.3}$$

where $\mu_{2,+}(\mathcal{M})$ denotes the set of nonnegative measures on \mathcal{M} whose support consists in at most two distinct points.

Proof. A direct calculation shows that

$$v_i(\theta, \mathbf{m}) = \sup_{x \in \mathbb{R}^2} \sup_{q \in \mu_+^x(\mathcal{M})} \left\{ a \left(x_1 \xi_1 + x_2 \theta - \int_{\mathcal{M}} \gamma(m) q(dm) \right) - \frac{a^2}{2} \|x\|^2 \right\}, \tag{B.4}$$

where

$$\mu_+^x(\mathcal{M}) = \left\{ q \in \mu_+(\mathcal{M}) : \int_{\mathcal{M}} \phi(m) q(dm) = x \right\}. \tag{B.5}$$

By Shapiro et al. (2014, Proposition 6.40), we have that the inner supremum remains the same if one optimizes over $\mu_+^x(\mathcal{M}) \cap \mu_{2,+}(\mathcal{M})$ rather than over $\mu_+^x(\mathcal{M})$, and the result follows. \square

Proof of Proposition 3.1. Fix a menu $\mathbf{m}_0 = (\gamma_0, b, \mathcal{M}_0)$ and consider an alternative menu of the form $\mathbf{m} = (1, \phi, \Theta)$ for some fund loading function $\phi : \Theta \rightarrow \mathbb{R}^2$. As a first step, we show that this fund loading function can be chosen in such a way that the investment strategy $\pi(\theta, \phi(\theta))$ delivers each investor the same utility as his best response to \mathbf{m}_0 and generates the same amount of management fees. By Lemma B.2, we have that given this menu each investor optimally allocates money to at most two funds. In order to construct the function $\phi(\theta)$, we therefore need to consider three mutually exclusive cases.

Case 0: If investors of type $\theta \in \Theta$ find it optimal to not invest in any of the proposed funds, then we know from Lemma B.1 that

$$\sup_{m \in \mathcal{M}_0} \{ b(m)^\top \xi(\theta) - \gamma_0(m) \} \leq 0. \tag{B.6}$$

Therefore, setting $\phi(\theta) = \frac{b(m)}{\gamma_0(m)}$ for some $m \in \mathcal{M}_0$ we get that $\pi(\theta, \phi(\theta)) = 0$ and the required properties follow.

Case 1: If the best response of investors of type $\theta \in \Theta$ is to allocate money to a single fund $m(\theta) \in \mathcal{M}_0$, then we have that

$$\begin{aligned} v_i(\theta, \mathbf{m}_0) &= u(\theta, r w_0 + q(\theta) \mathcal{R}(\gamma_0(m(\theta)), b(m(\theta)))) \\ &= \frac{1}{2 \|b(m(\theta))\|^2} (b(m(\theta))^\top \xi(\theta) - \gamma_0(m(\theta)))^2, \end{aligned} \tag{B.7}$$

where

$$\begin{aligned} q(\theta) &= \operatorname{argmax}_{q \in \mathbb{R}} \left\{ q(b(m(\theta))^\top \xi(\theta) - \gamma_0(m(\theta))) - \frac{a}{2} q^2 \|b(m(\theta))\|^2 \right\} \\ &= \frac{1}{a \|b(m(\theta))\|^2} (b(m(\theta))^\top \xi(\theta) - \gamma_0(m(\theta))). \end{aligned} \quad (\text{B.8})$$

Setting $\phi(\theta) := \frac{b(m(\theta))}{\gamma_0(m(\theta))}$ shows that we have $\gamma_0(m(\theta))q(\theta) = \pi(\theta, \phi(\theta))$ and the desired properties now follow by observing that

$$\pi(\theta, \phi(\theta))\mathcal{R}(1, \phi(\theta)) = q(\theta)\mathcal{R}(\gamma_0(m(\theta)), b(m(\theta))). \quad (\text{B.9})$$

Case 2: If the best response of investors of type $\theta \in \Theta$ is to allocate strictly positive amounts to a pair of funds $(m_1(\theta), m_2(\theta)) \in \mathcal{M}_0$ then

$$\begin{aligned} v_i(\theta, \mathbf{m}_0) &= u\left(\theta, rw_0 + \sum_{k=1}^2 q_k(\theta)\mathcal{R}(\gamma_0(m_k(\theta)), b(m_k(\theta)))\right) \\ &= \frac{a^2}{2} \left\| \sum_{k=1}^2 q_k(\theta)b(m_k(\theta)) \right\|^2, \end{aligned} \quad (\text{B.10})$$

where the vector

$$q(\theta) = \operatorname{argmax}_{q \in \mathbb{R}^2} u\left(\theta, rw_0 + \sum_{k=1}^2 q_k \mathcal{R}(\gamma_0(m_k(\theta)), b(m_k(\theta)))\right), \quad (\text{B.11})$$

satisfies the first-order conditions

$$b(m_k(\theta))^\top \xi(\theta) - \gamma_0(m_k(\theta)) = a \sum_{\ell=1}^2 q_\ell(\theta) (b(m_k(\theta))^\top b(m_\ell(\theta))). \quad (\text{B.12})$$

It follows that to satisfy the required properties for such types, we need to choose the fund loading function in such a way that

$$\begin{aligned} \sum_{k=1}^2 \gamma_0(m_k(\theta))q_k(\theta) &= \pi(\theta, \phi(\theta)), \\ \frac{a^2}{2} \left\| \sum_{k=1}^2 q_k(\theta)b(m_k(\theta)) \right\|^2 &= v_i(\theta, \mathbf{m}_0) = \frac{1}{2\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1), \end{aligned} \quad (\text{B.13})$$

and using the first-order conditions (B.12) shows that the unique solution to this system is explicitly given by

$$\phi(\theta) = \frac{q_1(\theta)b(m_1(\theta)) + q_2(\theta)b(m_2(\theta))}{q_1(\theta)\gamma_0(m_1(\theta)) + q_2(\theta)\gamma_0(m_2(\theta))}. \quad (\text{B.14})$$

To complete the proof, it now remains to show that the best response of an investor of type $\theta \in \Theta$ to the menu \mathbf{m} is indeed given by

$$q^*(\sigma; \theta, \mathbf{m}) = \mathbf{1}_{\{\theta \in \sigma\}} \pi(\theta, \phi(\theta)). \quad (\text{B.15})$$

As is easily seen, every fund in the menu \mathbf{m} is a linear combination of funds in \mathbf{m}_0 . Therefore, we have that $v_i(\theta, \mathbf{m}) \leq v_i(\theta, \mathbf{m}_0)$ for all $\theta \in \Theta$ and the desired result follows by observing that the definition of $\phi(\theta)$ guarantees that $v_i(\theta, \mathbf{m}_0) = u(\theta, w_1(q^*, \mathbf{m}))$. \square

Proof of Proposition 3.2. Fix a menu $\mathbf{m} = (1, \phi, \Theta)$ and for each $\theta \in \Theta$ denote by \mathcal{F}_θ the set of feasible directions from the measure defined in Equation (B.15). If the fund loading function is incentive compatible, then it follows from Lemma B.1 that

$$\sup_{\theta \in \Theta} \sup_{\nu \in \mathcal{F}_\theta} \int_{\Theta} \left\{ \phi(\theta')^\top \left(\xi(\theta) - a \int_{\Theta} \phi(\theta'') q^*(d\theta''; \theta, \mathbf{m}) \right) - 1 \right\} \nu(d\theta') \leq 0. \quad (\text{B.16})$$

Substituting the definition of the measure $q^*(\cdot; \theta, \mathbf{m})$ into the left hand side of this inequality, we obtain that

$$\sup_{\theta \in \Theta} \sup_{\nu \in \mathcal{F}_\theta} \int_{\Theta} \left\{ \phi(\theta')^\top \xi(\theta) - 1 - \frac{\phi(\theta')^\top \phi(\theta)}{\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1) \right\} \nu(\theta') \leq 0, \quad (\text{B.17})$$

and the validity of Equation (10) now follows by observing that the set \mathcal{F}_θ contains all nonnegative single point measures on Θ . Conversely, assume that the loading function $\phi : \Theta \rightarrow \mathbb{R}^2$ is such that Equation (10) holds, fix an arbitrary $\theta \in \Theta$, and let $\nu \in \mathcal{F}_\theta$ so that

$$\nu(\sigma) \geq -\mathbf{1}_{\{\theta \in \sigma\}} \pi(\theta, \phi(\theta)) / \delta, \quad (\text{B.18})$$

for some $\delta > 0$. Combining this property with Equation (10) shows that

$$\begin{aligned} & \int_{\Theta} \left\{ \phi(\theta')^\top \left(\xi(\theta) - a \int_{\Theta} \phi(\theta'') q^*(d\theta''; \theta, \mathbf{m}) \right) - 1 \right\} \nu(d\theta') \\ &= \int_{\Theta} \left\{ \phi(\theta')^\top \xi(\theta) - 1 - \frac{\phi(\theta')^\top \phi(\theta)}{\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1)_+ \right\} \nu(d\theta') \\ &= \{ \phi(\theta)^\top \xi(\theta) - 1 - (\phi(\theta)^\top \xi(\theta) - 1)_+ \} \nu(\{\theta\}) \\ & \quad + \int_{\Theta \setminus \{\theta\}} \left\{ \phi(\theta')^\top \xi(\theta) - 1 - \frac{\phi(\theta')^\top \phi(\theta)}{\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1)_+ \right\} \nu(d\theta') \\ & \leq -(1 - \phi(\theta)^\top \xi(\theta))_+ \nu(\{\theta\}) \leq (1 - \phi(\theta)^\top \xi(\theta))_+ \frac{(\phi(\theta)^\top \xi(\theta) - 1)_+}{\delta \|\phi(\theta)\|^2} = 0. \end{aligned} \quad (\text{B.19})$$

and the incentive compatibility of the fund loading function now follows from Lemma B.1 and the arbitrariness of the pair $(\theta, \nu) \in \Theta \times \mathcal{F}_\theta$. \square

B.2 | The relaxed problem

Proof of Lemma 3.3. Let $\phi \in \Phi$ and assume that

$$A := \{\theta \in \Theta : \pi(\theta, \phi(\theta)) > 0\} = \{\theta \in \Theta : \phi(\theta)^\top \xi(\theta) - 1 > 0\} \neq \emptyset, \tag{B.20}$$

for otherwise the statement is trivial. Since $\Phi \subseteq AC(\Theta; \mathbb{R}^2)$ and

$$\|\phi(\theta)\| \|\xi(\theta_H)\| \geq \|\phi(\theta)\| \|\xi(\theta)\| \geq \phi(\theta)^\top \xi > 1, \tag{B.21}$$

on the set A , we have that $v \in AC(\Theta; \mathbb{R})$. This implies that $v(\theta)$ is differentiable at almost every $\theta \in \Theta$ and a direct calculation shows that

$$\begin{aligned} \dot{v}(\theta) &= \frac{1}{\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1) (\dot{\phi}(\theta)^\top \xi(\theta) + \phi_2(\theta)) \\ &\quad - \frac{\dot{\phi}(\theta)^\top \phi(\theta)}{\|\phi(\theta)\|^4} (\phi(\theta)^\top \xi(\theta) - 1)^2, \end{aligned} \tag{B.22}$$

for almost every $\theta \in A$ and $\dot{v}(\theta) = 0$ for all $\theta \in A^c$. Equation (10) implies that for every fixed $\theta \in A$ the absolutely continuous function

$$\theta' \mapsto F_\theta(\theta') = (\phi(\theta')^\top \xi(\theta) - 1) - \frac{\phi(\theta')^\top \phi(\theta)}{\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1), \tag{B.23}$$

attains a maximum at the point $\theta' = \theta$. In particular, we have that

$$\left. \frac{\partial F_\theta(\theta')}{\partial \theta'} \right|_{\theta'=\theta} = \dot{\phi}(\theta)^\top \xi(\theta) - \frac{\dot{\phi}(\theta)^\top \phi(\theta)}{\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1) = 0 \tag{B.24}$$

for almost every $\theta \in A$. Substituting this expression into Equation (B.22) shows that

$$\dot{v}(\theta) = \frac{\phi_2(\theta)}{\|\phi(\theta)\|^2} (\phi(\theta)^\top \xi(\theta) - 1), \quad \text{a.e. } \theta \in A, \tag{B.25}$$

and combining this identity with the definition of $v(\theta)$ we obtain that

$$2v(\theta) \|\phi(\theta)\|^2 = (\phi(\theta)^\top \xi(\theta) - 1)^2, \tag{B.26}$$

$$\dot{v}(\theta) \|\phi(\theta)\|^2 = \phi_2(\theta) (\phi(\theta)^\top \xi(\theta) - 1), \tag{B.27}$$

for almost every $\theta \in A$. Solving this system gives

$$\phi_\pm(\theta) = \frac{\left(\sqrt{2v(\theta) - [\dot{v}(\theta)]^2}, \dot{v}(\theta) \right)^\top}{\theta \dot{v}(\theta) - 2v(\theta) \pm \xi \sqrt{2v(\theta) - [\dot{v}(\theta)]^2}}, \quad \text{a.e. } \theta \in A, \tag{B.28}$$

for almost every $\theta \in A$ and we claim that only $\phi_+(\theta)$ is consistent with the definition of the set A . Indeed, letting

$$p_{\pm}(\theta) := a\pi(\theta, \phi_{\pm}(\theta)) = \frac{\phi_{\pm}(\theta)^{\top} \xi(\theta) - 1}{\|\phi_{\pm}(\theta)\|^2}, \quad (\text{B.29})$$

and using Equation (B.25) in conjunction with Equation (B.28) and the definition of $v(\theta)$ we obtain that

$$\begin{aligned} p_{\pm}(\theta) &= xv(\theta) - 2v(\theta) \pm \xi \sqrt{2v(\theta) - [v(\theta)]^2} \\ &= x\phi_{\pm}(\theta)p_{\pm}(\theta) - \|\phi_{\pm}(\theta)\|^2 p_{\pm}(\theta)^2 \pm \xi \sqrt{\|\phi_{\pm}(\theta)\|^2 p_{\pm}(\theta)^2 - \phi_{2,\pm}(\theta)^2 p(\theta)^2} \\ &= x\phi_{\pm}(\theta)p_{\pm}(\theta) - \|\phi_{\pm}(\theta)\|^2 p_{\pm}(\theta)^2 \pm p_{\pm}(\theta)\phi_{1,\pm}(\theta)\xi \\ &= p_{\pm}(\theta) - (1 \mp 1)p_{\pm}(\theta)\phi_{1,\pm}(\theta)\xi, \quad \text{a.e. } \theta \in A. \end{aligned} \quad (\text{B.30})$$

This shows that

$$\mathbf{1}_{\{\theta \in A\}} \left(\phi(\theta)^{\top} - \frac{(\sqrt{2v(\theta) - [v(\theta)]^2}, v(\theta))}{\theta v(\theta) - 2v(\theta) + \xi \sqrt{2v(\theta) - [v(\theta)]^2}} \right) = 0, \quad (\text{B.31})$$

and it now remains to establish the validity of Equations (17) and (18). On the set A , these properties follow from Equations (B.29), (B.30), and the fact that we have

$$2v(\theta) - [v(\theta)]^2 = 2\phi_1(\theta)^2 v(\theta), \quad (\text{B.32})$$

as a result of Equations (B.26) and (B.27). On the set A^c we have that

$$\dot{v}(\theta) = v(\theta) = \frac{1}{\|\phi(\theta)\|^2} (\phi(\theta)^{\top} \xi(\theta) - 1)_+ = 0, \quad (\text{B.33})$$

and the desired result follows by observing that $F(\theta, 0, 0) = 0$ for all $\theta \in \Theta$. \square

The proof of Proposition 3.4 will be carried out through a series of lemmas. The first lemma establishes the uniqueness of the solution and shows that it attains the supremum in the relaxed problem.

Lemma B.3. *If $v^* \in C^2(\Theta; \mathbb{R})$ is a solution to the boundary value problem defined by Equations (22)–(24) then*

$$V = \int_{\Theta} F(\theta, v^*(\theta), \dot{v}^*(\theta)) d\theta. \quad (\text{B.34})$$

In particular, there can be at most one classical solution to the boundary value problem.

Proof. Let $v^* : \Theta \rightarrow \mathbb{R}$ be as in the statement. Since $v^* \in C^2(\Theta; \mathbb{R})$, we necessarily have that $v^* \in \mathcal{V}$ for otherwise

$$1 + \ddot{v}^*(\theta) = \frac{1}{v^*(\theta)} \left([v^*(\theta)]^2 + \frac{3}{2\xi} (2v^*(\theta) - [v^*(\theta)]^2)^{\frac{3}{2}} \right), \tag{B.35}$$

would not be real valued on Θ . Let now $v \in \mathcal{V}$ be another feasible function. Using Lemma B.4 below, we deduce that

$$\begin{aligned} & \int_{\Theta} (F(x, v(\theta), \dot{v}(\theta)) - F(x, v^*(\theta), \dot{v}^*(\theta))) d\theta \leq \Delta(v, v^*) \\ & := \int_{\Theta} ((v(\theta) - v^*(\theta))F_{v^*(\theta)}^*(\theta) + (\dot{v}(\theta) - \dot{v}^*(\theta))F_{\dot{v}^*(\theta)}^*(\theta)) d\theta, \end{aligned} \tag{B.36}$$

where we have set

$$F_k^*(\theta) := F_k(x, v^*(\theta), \dot{v}^*(\theta)), \quad k \in \{v^*(\theta), \dot{v}^*(\theta)\}. \tag{B.37}$$

Now, since $v^* \in C^2(\Theta; \mathbb{R})$ by assumption, we have that $\dot{v}^* \in C^1(\Theta; \mathbb{R})$ and we may thus integrate by parts to show that

$$\begin{aligned} \Delta(v, v^*) &= ((v - v^*)(\theta)F_p^*(\theta)) \Big|_{\theta=0}^{\theta_H} + \int_{\Theta} (v - v^*)(\theta) \left(F_v^*(\theta) - \frac{d}{d\theta} F_p^*(\theta) \right) d\theta \\ &= (v(\theta_H) - v^*(\theta_H))F_p^*(\theta_H) - (v(0) - v^*(0))F_p^*(0) = 0, \end{aligned} \tag{B.38}$$

where the last two equalities follow from the fact that v^* solves the Euler–Lagrange equation (20) subject to Equation (23). To complete the proof assume that $(v_i)_{i=1}^2 \in C^2(\Theta; \mathbb{R})$ are distinct solutions, and let

$$v^*(\theta) = \frac{1}{2} \sum_{i=1}^2 v_i(\theta). \tag{B.39}$$

By the first part of the proof we have that

$$V = \int_{\Theta} F(\theta, v_i(\theta), \dot{v}_i(\theta)) d\theta, \quad i \in \{1, 2\}, \tag{B.40}$$

and combining this identity with Equation (B.39) and Lemma B.4 we deduce that

$$V = \frac{1}{2} \sum_{i=1}^2 \int_{\Theta} F(\theta, v_i(\theta), \dot{v}_i(\theta)) d\theta < \int_{\Theta} F(x, v^*(\theta), \dot{v}^*(\theta)) d\theta. \tag{B.41}$$

Since $v^* \in \mathcal{V}$ this inequality contradicts the fact that the functions $(v_i)_{i=1}^2$ both attain the supremum in (\mathcal{R}) , and establishes the required uniqueness. \square

Lemma B.4. *Let*

$$\mathcal{O} := \{(v, p) \in \mathbb{R}^2 : v \neq 0 \text{ and } 2v - p^2 > 0\}. \tag{B.42}$$

The function $F(\theta, y, p)$ is strictly concave in $(v, p) \in \mathcal{O}$ for any fixed $\theta \in \Theta$.

Proof. A direct calculation shows that

$$\frac{\partial^2 F}{\partial v \partial p}(\theta, v, p) = \xi(2v - p^2)^{-\frac{3}{2}} \begin{bmatrix} -1 & p \\ p & -2v \end{bmatrix}. \tag{B.43}$$

The determinant and trace of this matrix are, respectively, strictly positive and strictly negative for all $(v, p) \in \mathcal{O}$. Therefore, its eigenvalues are strictly negative. \square

To prove the existence of a solution to the boundary value problem (22)–(24), we start by showing that for any initial condition q in an appropriate interval, the initial value problem given by Equation (22) subject to $\dot{v}(0) = v(0) - q = 0$ admits a unique classical solution. Then, we show that the initial condition q can be chosen in such a way as to satisfy the boundary condition (24).

Lemma B.5. *The initial value problem*

$$v(\theta)(1 + \ddot{v}(\theta)) - [\dot{v}(\theta)]^2 = \frac{3}{2\xi}(2v(\theta) - [\dot{v}(\theta)]^2)^{\frac{3}{2}}, \tag{B.44}$$

$$\dot{v}(0) = v(0) - q = 0, \tag{B.45}$$

admits a unique solution $v(\theta) = v(\theta; q)$ in $C^2(\mathbb{R}_+; \mathbb{R})$ for any $q > \frac{1}{18}\xi^2$. This solution is decreasing in q as well as strictly increasing and strictly convex in θ with

$$\inf_{\theta \geq 0} (\mathbf{1}_{\{q \in C_1\}} - \mathbf{1}_{\{q \in C_2\}})(1 - \ddot{v}(\theta)) \geq 0, \tag{B.46}$$

where we have set $C_1 := (\frac{1}{18}\xi^2, \frac{2}{9}\xi^2]$ and $C_2 := (2\xi^2/9, \infty)$.

Proof. Let q be fixed and write the initial value problem defined by Equations (B.44) and (B.45) as a system of first-order differential equations

$$0 = X'(\theta) - G(X(\theta)) = X(0) - (q, 0)^\top, \quad \theta \geq 0 \tag{B.47}$$

with the function

$$G(X) := \left(X_2, 1 + \frac{1}{\xi} \left(1 - \frac{X_2^2}{2X_1} \right) \left(3\sqrt{2X_1 - X_2^2} - 2\xi \right) \right)^\top. \tag{B.48}$$

Since $G \in C^1(\mathcal{O}; \mathbb{R}^2)$, it follows from Hirsch et al. (2013, p.387) that the initial value problem (B.47) admits a unique solution that is defined on $[0, \bar{\theta})$ for some $\bar{\theta} \leq \infty$. Before showing that this solution is actually defined on the whole positive real line, we start by establishing the other properties listed in the statement.

Letting the type $\theta \rightarrow 0$ in the differential equation and using the fact that $\dot{v}(0) = 0$ shows that we have

$$\xi(1 - \ddot{v}(0)) = 2\xi - \sqrt{18q}. \tag{B.49}$$

To proceed further, we distinguish two cases. If $q \in C_2$, then $\ddot{v}(0) > 1$, and we claim that the second derivative may reach one but never goes below. To see this, consider the function

$$b(\theta) := 2v(\theta) - [\dot{v}(\theta)]^2, \tag{B.50}$$

and assume that the solution is such that

$$\hat{\theta} := \inf \left\{ \theta \in [0, \bar{\theta}] : \ddot{v}(\theta) = 1 \right\} < \bar{\theta}, \tag{B.51}$$

for otherwise there is nothing to prove. Evaluating the differential equation (B.44) at the point $\hat{\theta}$ shows that

$$b(\hat{\theta})(2\xi - 3\sqrt{b(\hat{\theta})}) = 0, \tag{B.52}$$

and it follows that either $b(\hat{\theta}) = 0$ or $b(\hat{\theta}) = 4\xi^2/9$. Since $\ddot{v}(\theta) > 1$ on $[0, \hat{\theta})$ we have that $b(\theta)$ is strictly decreasing on that interval, and using this property in conjunction with the fact that $b(0) = 2q > 4\xi^2/9$, we deduce that $b(\hat{\theta}) = 4\xi^2/9$. This in turn implies that $v(\theta)$ solves the initial value problem

$$\begin{aligned} w(\theta)(1 + \ddot{w}(\theta)) &= [\dot{w}(\theta)]^2 + \frac{3}{2\xi}(2w(\theta) - [\dot{w}(\theta)]^2)^{\frac{3}{2}}, \\ w(\hat{\theta}) &= v(\hat{\theta}), \\ \dot{w}(\hat{\theta}) &= \left(2v(\hat{\theta}) - 4\xi^2/9\right)^{\frac{1}{2}}, \end{aligned} \tag{B.53}$$

on the interval $[\hat{\theta}, \bar{\theta})$ and, since the unique solution to this problem is

$$w(\theta) = v(\hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})^2 + (\theta - \hat{\theta})\left(2v(\hat{\theta}) - 4\xi^2/9\right)^{\frac{1}{2}}, \tag{B.54}$$

we conclude that $\ddot{v}(\theta) = \ddot{w}(\theta) = 1$ for all $\theta \in [\hat{\theta}, \bar{\theta})$. Because $\ddot{v}(\theta) \geq 1$ for all $\theta \in [0, \bar{\theta})$ we have that the solution is strictly convex and combining this property with the fact that $\dot{v}(0) = 0$ we deduce that it is strictly increasing.

Assume next that $q \in C_1$ so that $\ddot{v}(0) \in (0, 1]$. If $q = \frac{2}{9}\xi^2$ then a direct calculation shows that the unique solution to the initial value problem is

$$v(\theta) = \frac{2}{9}\xi^2 + \frac{1}{2}\theta^2, \tag{B.55}$$

and it follows that we have $\ddot{v}(\theta) = 1$ for all $\theta \in [0, \bar{\theta})$. Now assume that the initial condition $q < \frac{2}{9}\xi^2$, and denote by

$$\theta_0 := \inf \left\{ \theta \in [0, \bar{\theta}) : \ddot{v}(\theta) \geq 1 \right\}, \tag{B.56}$$

the first point at which the second derivative reaches one. Since $\ddot{v}(0) < 1$ we have that $\theta_1 > 0$. Assume that $\theta_1 < \bar{\theta}$. Since the solution is twice continuously differentiable on its domain of definition we have that $\ddot{v}(\theta_1) = 1$, and it thus follows from Equation (B.44) that the function $\ddot{v}(\theta)$ solves the differential equation

$$\dot{w}(\theta) = L(x, w(\theta)) := (1 - w(\theta)) \left(9\sqrt{b(\theta)} - 2\xi \right) \frac{\dot{v}(\theta)}{v(\theta)}, \tag{B.57}$$

$$w(\theta_1) = 1. \tag{B.58}$$

Since the functions $v(\theta)$ and $\dot{v}(\theta)$ are continuous in a neighborhood of θ_1 , we have that $L(x, w)$ is continuously differentiable in a neighborhood of $(\theta_1, 1)$ and it follows that there exists an $\varepsilon > 0$ such that Equations (B.57) and (B.58) admit a unique solution in $(\theta_1 - \varepsilon, \theta_1 + \varepsilon)$. Because $\ddot{v}(\theta)$ and the constant function $w(\theta) \equiv 1$ are both solutions to that differential equation, uniqueness implies that we must have $\ddot{v}(\theta) = 1$ in a left neighborhood of the point θ_1 . This contradicts the definition of θ_1 and thus shows that we have $\theta_1 = \infty$.

Since $\dot{v}(0) = 0$ and $\ddot{v}(0) > 0$ the strict increase of the solution will follow from its strict convexity. Assume towards a contradiction that the solution is not strictly convex on its domain of definition so that

$$\theta_0 := \inf \left\{ \theta \in [0, \bar{\theta}) : \ddot{v}(\theta) \leq 0 \right\} < \bar{\theta}. \tag{B.59}$$

By continuity we have that

$$0 = \ddot{v}(\theta_0) < \ddot{v}(\theta), \quad \theta \in [0, \theta_0), \tag{B.60}$$

and it follows that $v(\theta)$ and $\dot{v}(\theta)$ are strictly increasing on $[0, \theta_0)$. Differentiating both sides of Equation (B.44) shows that

$$\ddot{v}(\theta) = (1 - \ddot{v}(\theta)) \left(9\sqrt{b(\theta)} - 2\xi \right) \frac{\dot{v}(\theta)}{v(\theta)} \geq 0, \tag{B.61}$$

where the inequality follows from the nonnegativity and increase of $v(\theta)$, and the fact that $\ddot{v}(\theta) \leq 1$, which implies that

$$9\sqrt{b(\theta)} - 2\xi \geq 9\sqrt{b(0)} - 2\xi = 9\sqrt{2q} - 2\xi \geq \xi, \quad \theta \in [0, \bar{\theta}). \tag{B.62}$$

This shows that the function $\ddot{v}(\theta)$ is increasing on the interval $[0, \theta_0)$ and implies that we have $\ddot{v}(\theta_0) \geq \ddot{v}(0) > 0$, which contradicts Equation (B.60).

To complete the first part of the proof, it now remains to show that the solution is defined on the whole positive real line. Standard results on first-order differential systems (see e.g., Hirsch et al. (2013, p.398)) imply that this can fail to be the case only if the solution becomes unbounded or

reaches the boundary of \mathcal{O} . Assume first that $q \in C_1$. Using the fundamental theorem of calculus in conjunction with the fact that the solution is nondecreasing and such that $\ddot{v}(\theta) \leq 1$ gives

$$0 \leq v(\theta) = v(0) + \int_0^\theta \dot{v}(\theta)dx \leq v(0) + \theta = \theta, \tag{B.63}$$

and, therefore,

$$q = v(0) \leq v(\theta) = q + \int_0^\theta \dot{v}(\theta)dx = q + \int_0^\theta xdx = q + \frac{1}{2}\theta^2, \tag{B.64}$$

which shows that the solution cannot grow unbounded. On the other hand, because the solution is such that $\ddot{v}(\theta) \leq 1$, we have that the function $b(\theta)$ is nondecreasing, and it follows that

$$b(\theta) = 2v(\theta) - [\dot{v}(\theta)]^2 \geq b(0) = 2v(0) = 2q > 0, \tag{B.65}$$

which shows that the solution remains in \mathcal{O} . Assume next that the initial condition $q \in C_2$ and consider the function $b(\theta)$. Since $q \in C_2$, we know from the first part of the proof that this function is decreasing and such that $b(0) > 4\xi^2/9$. If the function $b(\theta)$ remains above $q^* = 4\xi^2/9$, then we have that the solution never reaches the boundary of the set \mathcal{O} . On the other hand, if the function $b(\theta)$ reaches q^* at some point $\theta^* \in [0, \bar{\theta})$, then it follows from Equation (B.44) that $\ddot{v}(\theta^*) = 1$ and the same arguments as in the first part of the proof then show that

$$v(\theta) = v(\theta^*) + \frac{1}{2}(\theta - \theta^*)^2 + (\theta - \theta^*)(2v(\theta^*) - 4\xi^2/9)^{\frac{1}{2}}, \tag{B.66}$$

for all $\theta \in [\theta^*, \bar{\theta})$. This in turn implies that $b(\theta) = 4\xi^2/9$ for all $\theta \in [\theta^*, \bar{\theta})$ and it follows that the solution to the initial value problem never reaches the boundary of \mathcal{O} . Finally, differentiating Equation (B.44) shows that

$$\ddot{v}(\theta) = (1 - \ddot{v}(\theta))\left(9\sqrt{b(\theta)} - 2\xi\right)\frac{\dot{v}(\theta)}{v(\theta)} \leq 0, \tag{B.67}$$

where the inequality follows from $\ddot{v}(\theta) \geq 1$ and $b(\theta) \geq 4\xi^2/9$. This implies that $\ddot{v}(\theta)$ is decreasing, and combining this with the strict increase of the solution we obtain that

$$v(0) \leq v(\theta) = q + \int_0^\theta (\theta - x)\ddot{v}(\theta)dx \leq q + \frac{1}{2}\ddot{v}(0)\theta^2, \tag{B.68}$$

and it follows that the solution cannot grow unbounded.

To complete the proof, it now remains to show that the solution is decreasing in the initial condition. Since the right hand side of Equation (B.47) belongs to $C^1(\mathcal{O}; \mathbb{R})$ we know from Hirsch et al. (2013, p.395) that the corresponding flow is continuous and the desired result will follow from the Kamke–Müller theorem (see, e.g., Müller (1927)) provided that the Jacobian matrix

$$J(\theta) := \nabla G(X(\theta)) = \left[\frac{\partial G_i}{\partial X_j}(X(\theta)) \right]_{i,j=1}^2, \tag{B.69}$$

is of Metzler type for all $\theta \geq 0$. A direct calculation shows that the off-diagonal terms of this matrix are explicitly given by $J_{12}(\theta) = 1$ and

$$J_{21}(\theta) = \frac{3(v(\theta) + [\dot{v}(\theta)]^2)\sqrt{2v(\theta) - [\dot{v}(\theta)]^2} - 2\xi[\dot{v}(\theta)]^2}{2\xi v(\theta)^2}. \tag{B.70}$$

Assume towards a contradiction that this function is not positive throughout the type space. Since $J_{21}(0) > 0$, this implies that there exists $\bar{\theta} > 0$ such that

$$J_{21}(\bar{\theta}) < 0 = J_{21}(\bar{\theta}). \tag{B.71}$$

The assumption that $J_{21}(\bar{\theta}) = 0$ implies that we have

$$\xi = \frac{3}{2[\dot{v}(\bar{\theta})^2]}(v(\bar{\theta}) + [\dot{v}(\bar{\theta})]^2)\sqrt{2v(\bar{\theta}) - [\dot{v}(\bar{\theta})]^2}. \tag{B.72}$$

Using this expression in conjunction with the fact that the function $v(\theta)$ solves Equation (B.44) the shows that we have

$$J_{21}(\bar{\theta}) = \frac{6\sqrt{2v(\bar{\theta}) - [\dot{v}(\bar{\theta})]^2}}{\dot{v}(\bar{\theta})(v(\bar{\theta}) + [\dot{v}(\bar{\theta})]^2)} \left\{ (v(\bar{\theta}) - [\dot{v}(\bar{\theta})]^2)^2 + v(\bar{\theta})[\dot{v}(\bar{\theta})]^2 \right\} \geq 0, \tag{B.73}$$

which contradicts Equation (B.71). □

Lemma B.6. *For any $\theta > 0$, there exists a unique $q = q(\theta) \in C_1$ such that*

$$\Lambda(\theta, q) := \theta - \frac{\xi \dot{v}(\theta; q)}{\sqrt{2v(\theta; q) - [\dot{v}(\theta; q)]^2}} = 0, \tag{B.74}$$

where $v(\theta; q)$ denotes the unique solution to (B.44) and (B.45). Furthermore, the function $q(\theta)$ is continuous, strictly decreasing, and such that $\lim_{\theta \rightarrow 0} q(\theta) = \frac{1}{8}\xi^2$.

Proof. Fix an arbitrary $\theta_H > 0$. Since the right hand side of Equation (B.47) belongs to $C^1(\mathcal{O}; \mathbb{R})$, we know from Hirsch et al. (2013, p.395) that the corresponding flow is continuous. Therefore, the function $q \mapsto \Lambda(\theta; q)$ is continuous on C_1 . A direct calculation shows that for $q \in \partial C_1$, this flow is explicitly given by

$$X\left(\theta; \frac{1}{18}\xi^2\right)^\top = \left(\frac{1}{18}\xi^2, 0\right), \tag{B.75}$$

$$X\left(\theta; \frac{2}{9}\xi^2\right)^\top = \left(\frac{2}{9}\xi^2 + \frac{1}{2}\theta^2, \theta\right). \tag{B.76}$$

Substituting these expressions in Equation (B.74) then shows that

$$\Lambda\left(\theta, \frac{2}{9}\xi^2\right) = -\frac{1}{2}\theta < 0 < \theta = \Lambda\left(\theta, \frac{1}{18}\xi^2\right), \tag{B.77}$$

and the existence of a solution now follows from the intermediate value theorem. To complete the first part of the proof assume that there exists $\theta_H > 0$ such that Equation (B.74) admits two solutions $q_1 \neq q_2$. Then the functions $v(\theta; q_1)$ and $v(\theta; q_2)$ both solve the boundary value problem associated with the upper end point $\theta_H > 0$ but differ in a right neighborhood of the origin. This contradicts the conclusion of Lemma B.3 and establishes the required uniqueness.

By the first part of the proof, we have that the solution mapping defines a function $q : (0, \infty) \rightarrow C$ such that $q(\theta) = v(\theta, q(\theta))$. Assume toward a contradiction that this function is not continuous so that there exist $\theta_0 > 0$ and $(\theta_{i,n})_{n=1}^\infty \subseteq (0, \infty)$ such that

$$\begin{aligned} \theta_0 &= \lim_{n \rightarrow \infty} \theta_{1,n} = \lim_{n \rightarrow \infty} \theta_{2,n}, \\ q_1 &:= \lim_{n \rightarrow \infty} q(\theta_{1,n}) \neq \lim_{n \rightarrow \infty} q(\theta_{2,n}) := q_2. \end{aligned} \tag{B.78}$$

The continuity of $(v(\theta; q), \dot{v}(\theta; q))$ and the definition of $q(\theta)$ then imply that

$$0 = \lim_{n \rightarrow \infty} \left| \Lambda(\theta_{i,n}, q(\theta_{i,n})) \right| = |\Lambda(\theta_0, q_i)|, \tag{B.79}$$

and it follows that $q_i \in \text{int}(C_1)$ for otherwise Equation (B.77) would imply that the term on the right hand side is strictly positive. This contradicts the fact that the solution to Equation (B.74) is unique in C_1 for every $\theta > 0$, and establishes the required continuity.

Now assume that the solution mapping is not strictly monotone. By continuity this implies that there exist $\theta_1, \theta_2 > 0$ such that $\theta_1 \neq \theta_2$ and $q(\theta_1) = q(\theta_2) := q^* \in C_1$. The definition of the solution mapping then implies that

$$\Lambda(\theta_1, q^*) = \Lambda(\theta_2, q^*), \tag{B.80}$$

which contradicts Lemma B.7 below. Next, we claim that the solution mapping is such that $q(\theta) < \frac{1}{8}\xi^2$ for all $\theta > 0$. Indeed, if we had that $q(\theta_0) \geq \frac{1}{8}\xi^2$ for some $\theta_0 > 0$ then

$$\theta \mapsto \Lambda(\theta, (q(\theta_0))) = 0, \tag{B.81}$$

would admit a solution given by $\theta = \theta_0 > 0$ and this would contradict Lemma B.7.

The above results show that the solution mapping is continuous, monotone and bounded on $(0, \infty)$. Therefore, the limit $q(0) := \lim_{x \rightarrow 0} q(\theta)$ exists and the proof will be complete once we show that equals the constant in the statement. Since $\dot{v}(0; q) = 0$ for all $q \in C_1$ we have that

$$\lim_{\theta \rightarrow 0} \frac{\dot{v}(\theta, q(\theta))}{\theta} = \dot{v}(0; q(0)). \tag{B.82}$$

Using this identity in conjunction with Equation (B.49) and the definition of the solution mapping we obtain that

$$0 = \lim_{\theta \rightarrow 0} \frac{\Lambda(\theta, q(\theta))}{\theta} = 1 - \frac{\xi \dot{v}(0; q(0))}{\sqrt{2q(0)}} = 1 + \frac{\xi}{\sqrt{2q(0)}} \left(1 - \frac{\sqrt{18q(0)}}{\xi} \right), \tag{B.83}$$

and solving for $q(0)$ gives the desired result. Knowing that the solution mapping is strictly monotone and such that $q(\theta) \leq q(0)$ for all $\theta \geq 0$, we then deduce that it is strictly decreasing and the proof is complete. \square

Lemma B.7. *For $q \in C_1$ the equation $\Lambda(\theta, q) = 0$ admits a solution $\theta > 0$ only if $q \leq \frac{1}{8}\xi^2$, and in this case there is at most one solution.*

Proof. A direct calculation using Equation (B.44) shows that

$$\frac{\partial \Lambda}{\partial \theta}(\theta, q) = \frac{\xi}{\sqrt{2v(\theta; q) - [\dot{v}(\theta; q)]^2}} - 2. \tag{B.84}$$

Since $\ddot{v}(\theta; q) \leq 1$ for all $(\theta, q) \in \mathbb{R}_+ \times C_1$, by Lemma B.5 we have that this derivative is nonincreasing as a function of θ , and it follows that the function $\theta \mapsto \Lambda(\theta, q)$ is concave. If $q > \frac{1}{8}\xi^2$, then this concavity implies that

$$\frac{\partial \Lambda}{\partial \theta}(\theta, q) \leq \frac{\partial \Lambda}{\partial \theta}(0, q) = \frac{\xi}{\sqrt{2q}} - 2 < 0, \tag{B.85}$$

and it follows that the only solution to $\Lambda(\theta, q) = 0$ is given by $\theta = 0$. On the other hand, if $q \leq \frac{1}{8}\xi^2$, then $\frac{\partial \Lambda}{\partial \theta}(0, z) \geq 0$, and the concavity of the function $\theta \mapsto \Lambda(\theta, q)$ implies that there can be at most one $\theta > 0$ such that $\Lambda(\theta, q) = 0$. \square

Proof of Proposition 3.4. By construction $v^*(\theta) = v(\theta; q(\theta_H))$ belongs to $C^2(\Theta; \mathbb{R})$ and solves the boundary value problem. Therefore, it follows from Lemma B.3 that this function is the unique such solution and that it attains the supremum in the relaxed problem. Furthermore, since $\theta_H > 0$, we know from Lemma B.6 that $q(\theta_H) > \frac{1}{8}\xi^2$, and it thus follows from Lemma B.5 that $v^*(\theta)$ is strictly increasing and strictly convex. \square

B.3 | The optimal fund menu

Proof of Theorem 3.5. Let us start by establishing Equation (25). Since $v^*(0) = q(\theta_H) \in C_1$ by Lemma B.6, we know from Lemma B.5 that

$$\inf_{\theta \in \Theta} (1 - \dot{v}^*(\theta)) \geq 0. \tag{B.86}$$

Using this and the fundamental theorem of calculus then shows that we have

$$2v^*(\theta) - [\dot{v}^*(\theta)]^2 = 2q(\theta_H) + \int_0^\theta 2\dot{v}^*(x)(1 - \dot{v}^*(x))dx > \frac{1}{9}\xi^2. \tag{B.87}$$

As a result, Equation (25) will follow once we show that

$$c(\theta) := \frac{aF(\theta, v^*(\theta), \dot{v}^*(\theta))}{\sqrt{2v^*(\theta) - [\dot{v}^*(\theta)]^2}} = \xi + \frac{\theta \dot{v}^*(\theta) - 2v^*(\theta)}{\sqrt{2v^*(\theta) - [\dot{v}^*(\theta)]^2}} > 0, \tag{B.88}$$

for all $\theta \in \Theta$. Since $v^*(0) = q(\theta_H) \in C_1$ by Lemma B.6, we have that $c(0) \geq \xi/3$ is strictly positive, and it is therefore sufficient to show that $c(\theta)$ is nondecreasing. Differentiating Equation (B.88), we find that

$$\dot{c}(\theta) = \frac{(\theta - \dot{v}^*(\theta))(2v^*(\theta)\ddot{v}^*(\theta) - [\dot{v}^*(\theta)]^2)}{(2v^*(\theta) - [\dot{v}^*(\theta)]^2)^{\frac{3}{2}}}. \tag{B.89}$$

Using Equation (B.86), we deduce that

$$\theta - \dot{v}^*(\theta) = \int_0^\theta (1 - \ddot{v}^*(x))dx \geq 0. \tag{B.90}$$

On the other hand, using the fact that $v^*(0) = q(\theta_H) \in C_1$ in conjunction with the fundamental theorem of calculus, Equations (B.49), and (B.61) we obtain that

$$\begin{aligned} \ell(\theta) &:= 2v^*(\theta)\ddot{v}^*(\theta) - [\dot{v}^*(\theta)]^2 = \ell(0) + \int_0^\theta 2v^*(x)\ddot{v}^*(x)dx \\ &\geq \ell(0) = 2v^*(0) \left(\sqrt{\frac{18v^*(0)}{\xi^2}} - 1 \right) \geq 0, \end{aligned} \tag{B.91}$$

and the desired result now follows from Equation (B.87).

Let us now turn to the second part of the statement. Since $v^* \in C^2(\Theta; \mathbb{R})$, we have that $\phi^* \in C^1(\Theta; \mathbb{R}^2)$, and the feasibility of ϕ^* will follow once we show that

$$L(\theta, \theta') := (\phi^*(\theta')^\top \xi(\theta) - 1) - \frac{\phi^*(\theta')^\top \phi^*(\theta)}{\|\phi^*(\theta)\|^2} (\phi^*(\theta)^\top \xi(\theta) - 1)_+, \tag{B.92}$$

is nonpositive for all $(\theta, \theta') \in \Theta^2$. Substituting the definition of $\phi^*(\theta)$ into the left hand side, and using the fact that

$$\frac{1}{\|\phi^*(\theta)\|^2} (\phi^*(\theta)^\top \xi(\theta) - 1)_+ = aF(\theta, v^*(\theta), \dot{v}^*(\theta)) > 0, \tag{B.93}$$

as a result of Equation (25), we obtain that

$$L(\theta, \theta') = -\frac{2v^*(\theta)(g(\theta)g(\theta') - 2v^*(\theta') + (\theta' - \theta + \dot{v}^*(\theta))\dot{v}^*(\theta'))}{a^3F(\theta', v^*(\theta'), \dot{v}^*(\theta'))F(\theta, v^*(\theta), \dot{v}^*(\theta))^2}, \tag{B.94}$$

with

$$g^*(\theta) := \sqrt{2v^*(\theta) - [\dot{v}^*(\theta)]^2}. \tag{B.95}$$

Since the functions $v^*(\theta)$ and $F(\theta, v^*(\theta), \dot{v}^*(\theta))$ are both strictly positive on Θ , it is sufficient to show that

$$g^*(\theta) \geq h(\theta; \theta') := \frac{2v^*(\theta') + (\theta - \theta' - \dot{v}^*(\theta))\dot{v}^*(\theta')}{g^*(\theta')}, \quad (\theta, \theta') \in \Theta^2, \tag{B.96}$$

but, because $h(\theta; \theta) = g(\theta)$ for all types, it is actually enough to show that the map $\theta' \mapsto h(\theta; \theta')$ is increasing on $[0, \theta]$ and decreasing on $[\theta, \theta_H]$. A direct calculation shows that

$$[g^*(\theta')]^3 \frac{\partial h}{\partial \theta'}(\theta; \theta') = (\theta - v^*(\theta) - \theta' + v\theta') \ell(\theta'), \quad (\text{B.97})$$

and Equation (B.91) implies that the sign of this derivative is the same as that of

$$(\theta - v^*(\theta)) - (\theta' - v^*(\theta')). \quad (\text{B.98})$$

Due to Equation (B.86), we have that the function $\theta - v^*(\theta)$ is nondecreasing. Therefore, the sign of the above expression is the same as that of the difference $\theta - \theta'$. \square

Lemma B.8. *The map $\theta \mapsto F(\theta, v^*(\theta), v^*(\theta))$ is nondecreasing.*

Proof. A direct calculation shows that

$$\frac{dF}{d\theta}(\theta, v^*(\theta), v^*(\theta)) = \left(\theta - \frac{v^*(\theta)}{g^*(\theta)} \xi \right) \dot{v}^*(\theta) + \left(\frac{\xi}{g^*(\theta)} - 1 \right) \dot{v}^*(\theta). \quad (\text{B.99})$$

Since the function $v^*(\theta)$ is increasing and convex by Lemma B.5, we only need to show that the bracketed terms are nonnegative. Consider the first term. Since

$$\theta - \frac{v^*(\theta)}{g^*(\theta)} \xi = \Lambda(\theta; v^*(\theta)), \quad (\text{B.100})$$

we know from the proof of Lemma B.7 that this term is concave in θ , equal to zero on the boundary of the type space, and such that

$$\left. \frac{d}{d\theta} \left(\theta - \frac{v^*(\theta)}{g^*(\theta)} \xi \right) \right|_{\theta=0} > 0. \quad (\text{B.101})$$

This implies that this term is nonnegative throughout Θ . On the other hand, a direct calculation using the definition of the function $g^*(\theta)$ shows that

$$\frac{d}{d\theta} \left(\frac{\xi}{g^*(\theta)} - 1 \right) = - \frac{\dot{g}^*(\theta) \xi}{[g^*(\theta)]^2} = - \frac{v^*(\theta)(1 - v^*(\theta)) \xi}{[g^*(\theta)]^{5/2}}. \quad (\text{B.102})$$

Since $v^*(0) \in C_1$, we know from Lemma B.5 that $\dot{v}^*(\theta) \leq 1$ and $v^*(\theta) \geq 0$. Therefore, the above expression is negative and the desired result now follows by observing that

$$\frac{g^*(\theta_H)}{\xi} = \frac{v^*(\theta_H)}{\theta_H} = \int_0^{\theta_H} v^*(\theta) \frac{d\theta}{\theta_H} \geq 1, \quad (\text{B.103})$$

due to Equation (24), the fundamental theorem of calculus and Lemma B.5. \square

Proof of Proposition 4.2. A direct calculation using the fact that $v^*(\theta)$ solves Equation (22) subject to Equations (23) and (24) shows that we have

$$\begin{aligned}\Delta(0) &= \Delta(\theta_H) = 0, \\ \dot{\Delta}(\theta) &= \frac{1}{g^*(\theta)\xi} \left(\frac{\xi}{2} - g^*(\theta) \right),\end{aligned}\tag{B.104}$$

and therefore $\dot{\Delta}(0) > 0$ since $v^*(0) < \frac{1}{8}\xi^2$ by Lemma B.6. In view Lemma B.5, we have that $g^*(\theta)$ is nonnegative and increasing. Therefore, the derivative $\dot{\Delta}(\theta)$ only changes sign once, and the desired result now follows from the fact that the function is equal to zero on the boundary of the type space. \square

Proof of Proposition 4.4. To establish the existence of a constant θ_2 with the required property we need to show that the function

$$\ell(\theta) := a(\pi(\theta, \phi^*(\theta))\phi_1^*(\theta) - q^\circ(\theta)\phi_1^\circ(\theta)) = g^*(\theta) - \frac{\xi}{2},\tag{B.105}$$

is first negative then positive. Since $\dot{v}^*(\theta) \leq 1$ and $v^*(\theta) \geq 0$, by Lemma B.5, we have that the function $\ell(\theta)$ is increasing, and it is thus sufficient to show that it crosses the horizontal axis. Consider the function

$$\Delta(\theta) := \frac{\pi(\theta, \phi^*(\theta))\phi_2^*(\theta)}{\pi(\theta, \phi^*(\theta))\phi_1^*(\theta)} - \frac{q^\circ(\theta)\phi_2^\circ(\theta)}{q^\circ(\theta)\phi_1^\circ(\theta)} = \frac{\dot{v}^*(\theta)}{g^*(\theta)} - \frac{\theta}{\xi}.\tag{B.106}$$

As shown in the proof of Proposition 4.2, we have that the derivative of this function changes sign only once and the desired result now follows by observing that

$$\dot{\Delta}(\theta) = \frac{\ell(\theta)}{g^*(\theta)\xi}.\tag{B.107}$$

To show the existence of a constant θ_1 with the required property, consider the function

$$k(\theta) := a(\pi(\theta, \phi^*(\theta))\phi_2^*(\theta) - q^\circ(\theta)\phi_2^\circ(\theta)) = \dot{v}^*(\theta) - \frac{\theta}{2}.\tag{B.108}$$

Combining Lemmas B.5 and B.6, we deduce that this function is increasing, convex, and such that

$$\dot{k}(0) = \dot{v}^*(0) - \frac{1}{2} < 0 = k(0).\tag{B.109}$$

Therefore, the function $k(\theta)$ crosses the horizontal axis at most once and the existence of a constant θ_1 with the required property now follows by observing that, due to Equation (24), the increase of the function of $g^*(\theta)$ and the definition of θ_2 , we have

$$k(\theta_H) = \dot{v}^*(\theta_H) - \frac{\theta_H}{2} = \frac{\theta_H}{\xi} \left(g^*(\theta_H) - \frac{\xi}{2} \right) \geq \frac{\theta_H}{\xi} \left(g^*(\theta_2) - \frac{\xi}{2} \right) = 0.\tag{B.110}$$

Let us now show that the constants θ_1 and θ_2 are such that $\theta_1 \leq \theta_2$. Since

$$\frac{\pi(\theta, \phi^*(\theta))\phi_2^*(\theta)}{\pi(\theta, \phi^*(\theta))\phi_1^*(\theta)} \leq \frac{q^\circ(\theta)\phi_2^\circ(\theta)}{q^\circ(\theta)\phi_1^\circ(\theta)}, \tag{B.111}$$

for all $\theta \in \Theta$, by Proposition 4.2 it follows from the definition of θ_1 that we have

$$\begin{aligned} \pi(\theta_1, \phi^*(\theta_1))\phi_2^*(\theta_1) &\leq \pi(\theta_1, \phi^*(\theta_1))\phi_1^*(\theta_1) \frac{q_2(\theta_1)\phi_2^\circ(\theta_1)}{q_2(\theta_1)\phi_1^\circ(\theta_1)} \\ &= \pi(\theta_1, \phi^*(\theta_1))\phi_1^*(\theta_1). \end{aligned} \tag{B.112}$$

Therefore, $k(\theta_1) \leq 0$ and the desired conclusion now follows from the first part of the proof. The remaining claims in the statement follow from Lemma B.9 below. \square

Lemma B.9. *There exists $\bar{\theta} \in [\theta_1, \theta_2]$ such that $v^*(\theta) \leq v^\circ(\theta)$ if and only if $\theta \leq \bar{\theta}$.*

Proof. Consider the function defined by

$$m(\theta) := v^*(\theta) - v^\circ(\theta) = v^*(\theta) - \frac{1}{8}\|\xi(\theta)\|^2. \tag{B.113}$$

Since $v^*(0) < \xi^2/8$ by Lemma B.6, we have that $m(0) < 0$. On the other hand, the result of Proposition 4.4 implies that we have

$$\begin{aligned} m(\theta_H) &= v^*(\theta_H) - \frac{1}{8}\|\xi(\theta_H)\|^2 \\ &= \frac{1}{2}[g^*(\theta_H)]^2 + \frac{1}{2}[v^*(\theta_H)]^2 - \frac{1}{8}\|\xi(\theta_H)\|^2 \\ &\geq \frac{1}{2}[\xi/2]^2 + \frac{1}{2}[\theta_H/2]^2 - \frac{1}{8}\|\xi(\theta_H)\|^2 = 0, \end{aligned} \tag{B.114}$$

and it follows from the intermediate value theorem that there exist $\bar{\theta} \in \Theta$ such that $m(\bar{\theta}) = 0$. To complete the proof, it is now sufficient to show that this point is unique and lies between θ_1 and θ_2 . A direct calculation gives

$$0 = \dot{m}(\bar{\theta}) - \left(\dot{v}(\bar{\theta}) - \frac{1}{4}\bar{\theta} \right) = \dot{m}(\bar{\theta}) - \left(\ddot{v}(\bar{\theta}) - \frac{1}{4} \right), \tag{B.115}$$

and, as shown in the proof of Lemma B.5, we have that $\ddot{v}^*(\theta)$ is nondecreasing. It follows that two cases may occur. If $\dot{m}(0) > 0$, then the function $m(\theta)$ is convex and therefore increasing, which implies that it can cross the horizontal axis at most once. On the contrary, if $\dot{m}(0) \leq 0$, then $\dot{m}(\theta)$ changes sign at most once and the existence of unique crossing point follows. Finally, using Proposition 4.4, we obtain that

$$\begin{aligned} v^*(\theta_1) &= \frac{1}{2}[v^*(\theta_1)]^2 + \frac{1}{2}[g^*(\theta_1)]^2 = \frac{1}{2}[v^*(\theta_1)]^2 + \frac{1}{8}\xi^2 \\ &\leq \frac{1}{2}[\theta_1/2]^2 + \frac{1}{8}\xi^2 = \frac{1}{8}\|\xi(\theta_1)\|^2, \end{aligned}$$

$$\begin{aligned} v^*(\theta_2) &= \frac{1}{2}[v^*(\theta_2)]^2 + \frac{1}{2}[g^*(\theta_2)]^2 = \frac{1}{2}[\theta_2/2]^2 + \frac{1}{2}[g^*(\theta_2)]^2 \\ &\geq \frac{1}{2}[\theta_2/2]^2 + \frac{1}{2}[\xi/2]^2 = \frac{1}{8}\|\xi(\theta_2)\|^2, \end{aligned} \quad (\text{B.116})$$

and the desired result now follows from the first part of the proof. \square

Lemma B.10. *The functions $\phi_1^*(\theta)$ and $\phi_2^*(\theta)$ are, respectively, decreasing and increasing with respect to the investor type.*

Proof. A direct calculation shows that we have

$$\dot{\phi}_1^*(\theta) = \frac{(\theta - v^*(\theta))g^*(\theta)}{\xi F(\theta, v^*(\theta), v^*(\theta))^2} ([v^*(\theta)]^2 - 2v^*(\theta)v^{**}(\theta)). \quad (\text{B.117})$$

Since $v^{**}(\theta) \leq 1$, by Lemma B.5 and $v^*(0) = 0$ we have that $\theta - v^*(\theta) \geq 0$. Therefore, the sign of the above derivative depends on the sign of the bracketed term on the right hand side and we know from the proof of Theorem 3.5 that this term is negative throughout the type space. Similarly, a direct calculation gives

$$\dot{\phi}_2^*(\theta) = \frac{(g^*(\theta) - \xi)}{g^*(\theta)F(\theta, v^*(\theta), v^*(\theta))^2} ([v^*(\theta)]^2 - 2v^*(\theta)v^{**}(\theta)), \quad (\text{B.118})$$

and the same argument as in the first part of the proof show that the sign of this quantity depends on that of the function $\xi - g^*(\theta)$. Because $v^{**}(\theta) \leq 1$, we have that this function is increasing and the desired result now follows by observing that

$$g^*(\theta_H) = (\xi/\theta_H)v^*(\theta_H) \leq \xi, \quad (\text{B.119})$$

as a result of Equation (24) and the fact that $v^*(\theta) \leq \theta$ for all $\theta \in \Theta$. \square

Proof of Proposition 3.6. This directly follows by combining Lemmas B.8 and B.10. \square

Proof of Proposition 4.1. By Theorem 3.5 we have that

$$M = \sup_{v \in \Phi} \int_0^{\theta_H} F(\theta, v(\theta), v(\theta)) \frac{d\theta}{\theta_H} = \int_0^{\theta_H} F(\theta, v^*(\theta), v^*(\theta)) \frac{d\theta}{\theta_H}. \quad (\text{B.120})$$

Therefore, the envelope theorem implies that

$$\frac{dM}{d\xi} = \int_0^{\theta_H} \frac{dF}{d\xi}(\theta, v^*(\theta), v^*(\theta)) \frac{d\theta}{\theta_H} = \int_0^{\theta_H} g^*(\theta) \frac{d\theta}{\theta_H}, \quad (\text{B.121})$$

and the required monotonicity in ξ follows from the nonnegativity of $g^*(\theta)$. Similarly, an application of the envelope theorem shows that

$$\frac{dM}{d\theta_H} = \frac{1}{\theta_H^2} \int_0^{\theta_H} (F(\theta_H, v^*(\theta_H), v^*(\theta_H)) - F(\theta, v^*(\theta), v^*(\theta))) d\theta, \quad (\text{B.122})$$

and the desired monotonicity follows from Lemma B.8. \square

Proof of Lemma 4.5. By Proposition 3.1, we have that there exists an incentive compatible fund loading function ϕ_U that implements the same indirect utilities as the unbundling solution. Therefore $M_0 \leq M$ and the result will follow once we show that the inequality is strict. Let

$$v_U(\theta) = \frac{1}{2\|\phi_U(\theta)\|^2} (\phi_U(\theta)^\top \xi(\theta) - 1)_+^2 = \frac{1}{8}\xi^2 + \frac{1}{2}\left(\theta - \frac{\theta_H}{3}\right)_+^2, \quad (\text{B.123})$$

denote the indirect utility that investors derive when offered ϕ_U . By Lemma 3.3, we have that $v_U \in \mathcal{V}$ and it thus follows from Theorem 3.5 that

$$M_0 = \int_{\Theta} F(\theta, v_U(\theta), \dot{v}_U(\theta)) \frac{d\theta}{\theta_H} \leq M = \sup_{v \in \mathcal{V}} \int_{\Theta} F(\theta, v(\theta), \dot{v}(\theta)) \frac{d\theta}{\theta_H}. \quad (\text{B.124})$$

Since the supremum on the right is uniquely attained by the function v^* , it now suffices to show that the functions v^* and v_U differ over an open subset of the type space and this property follows by observing that over the interval $(0, \theta_H/3)$, the function v_U is constant whereas the function v^* is strictly increasing. \square

B.4 | Performance fees

If the manager is allowed to charge performance fees as in Equation (8), then the budget constraint of an investor is given by

$$w_1 = rw_0 + \int_{\mathcal{M}} \left((\delta(m)b(m) + (1 - \delta(m))\phi(m))^\top \varepsilon - \gamma(m) \right) q(dm), \quad (\text{B.125})$$

where $q \in \mu_+(\mathcal{M})$ is a positive measure. Comparing this expression to Equation (4) reveals that from the point of view of an investor, the *extended* menu

$$\{(\gamma(m), \delta(m), b(m), \phi(m)) : m \in \mathcal{M}\}, \quad (\text{B.126})$$

is payoff equivalent to the standard menu with loading function $\delta(m)b(m) + (1 - \delta(m))\phi(m)$ and fee rate $\gamma(m)$. Using this property and proceeding as in Section 3 shows that the manager can focus on extended menus of the form

$$\left\{ \left(1, \delta(\theta), b(\theta), \frac{\phi(\theta) - \delta(\theta)b(\theta)}{1 - \delta(\theta)} \right) : \theta \in \Theta \right\}, \quad (\text{B.127})$$

where $(\delta, b) : \Theta \rightarrow (\mathbb{R}_+, \mathbb{R}^2)$ are arbitrary functions and $\phi \in \Phi_0$ is an incentive compatible loading function in the sense of Proposition 3.1. A direct calculation then shows that the amount of fees generated by such a menu is *random* and given by

$$\int_{\Theta} \pi(\theta, \phi(\theta)) \left[1 + \frac{\delta(\theta)}{1 - \delta(\theta)} (\phi(\theta) - b(\theta))^\top \varepsilon \right] \frac{d\theta}{\theta_H}. \quad (\text{B.128})$$

Since the manager represents an investment firm, it is natural to assume that he has access to a wider financial market in which any risk can be hedged. This implies that the manager should use an equivalent martingale measure to value the present values of the fees and since ε has expectation zero under any such measure we conclude that the optimal menu can be implemented without using performance fees

APPENDIX C: PROOF OF THE RESULTS IN SECTION 5

C.1 | Incentive compatible menus

Proposition C.1. *Assume that investors can directly access the familiar asset. Then, given any fund menu $\bar{\mathbf{m}}$ there exists a fund menu $\mathbf{m} = (\gamma, \phi, \mathcal{M})$ such that*

- (1) $\mathcal{M} = \Theta$,
- (2) $\gamma(\theta) = 1$ for all $\theta \in \Theta$,
- (3) $q^*(\sigma; \theta, \mathbf{m}) = \mathbf{1}_{\{\theta \in \sigma\}} \pi(\theta, \phi(\theta))$ and $n^*(\theta; \mathbf{m}) = 0$ for all $\theta \in \Theta$ and $\sigma \subseteq \Theta$,
- (4) $v_m(\mathbf{m}) = v_m(\bar{\mathbf{m}})$, and
- (5) $v_i(\theta, \mathbf{m}) = v_i(x, \bar{\mathbf{m}})$ for all $\theta \in \Theta$.

Proof of Propositions C.1 and 5.2. The proofs of these propositions are similar to those of Propositions 3.1 and 3.2. We omit the details. \square

Proof of Lemma 5.3. The proof is similar to that of Lemma 3.3. We omit the details. \square

C.2 | The relaxed problem

Proof of Lemma 5.1. Since $\Phi_1 \subseteq \Phi$, we have that $M_1 \leq M$ and it follows that it is sufficient to show that v^* satisfies Equation (46). The definition of the function v^* and the results of Lemmas B.5 and B.6 imply that $\dot{v}^*(\theta) \leq 1$ for all $\theta \in \Theta$. Therefore, the function $b^*(\theta) := 2v^*(\theta) - [\dot{v}^*(\theta)]^2$ is non-decreasing, and the desired result follows by noting that under the assumption of the statement, we have $b^*(0) \geq (\xi - \gamma_1)^2$. \square

To derive a set of optimality conditions for the relaxed problem (\mathcal{R}_1) , we consider the Lagrangian objective defined by

$$\int_{\Theta} H^\lambda(\theta, v(\theta), \dot{v}(\theta)) d\theta := \int_{\Theta} F(\theta, v(\theta), \dot{v}(\theta)) d\theta + \int_{\Theta} \lambda(\theta) c(v(\theta), \dot{v}(\theta)) d\theta, \quad (\text{C.1})$$

where the function

$$c(v(\theta), \dot{v}(\theta)) := 2v(\theta) - [\dot{v}(\theta)]^2 - (\xi - \gamma_1)^2, \quad (\text{C.2})$$

returns the value of the constraint (46) and $\lambda : \Theta \rightarrow \mathbb{R}_+$ is a Lagrange multiplier that enforces this constraint at each point of the type space.

The next lemma provides a set of sufficient optimality conditions for the relaxed problem (\mathcal{R}_1) . To state the result, denote by $AC_p^*(\Theta; \mathbb{R})$ the set of piecewise absolutely continuous functions that are right continuous at zero, left continuous at θ_H , and have at most finitely many jump discontinuities.

Lemma C.2. *Let $(v, \lambda) \in \mathcal{V}_1 \times AC_p^*(\Theta; \mathbb{R}_+)$ be such that $\dot{v} \in AC(\Theta; \mathbb{R})$ and denote by \mathcal{C} the set of points where the function λ is continuous. If*

$$\left(H_{v(\theta)}^\lambda - \frac{d}{d\theta} H_{\dot{v}(\theta)}^\lambda \right) (\theta, v(\theta), \dot{v}(\theta)) = 0, \quad \theta \in \mathcal{C}, \tag{C.3}$$

$$H_{\dot{v}(\theta)}^\lambda (\theta, v(\theta), \dot{v}(\theta)) = 0, \quad \theta \in \{0, \theta_H\}, \tag{C.4}$$

$$\lambda(\theta)c(v(\theta), \dot{v}(\theta)) = 0, \quad \theta \in \Theta, \tag{C.5}$$

and $H_{\dot{v}(\theta)}^\lambda (\theta, v(\theta), \dot{v}(\theta))$ is continuous, then v attains the supremum in (\mathcal{R}_1) .

The above conditions can be interpreted as follows. The first condition requires that the Euler–Lagrange equation associated with the optimization of the Lagrangian holds at all points of continuity of the multiplier. The second condition imposes the boundary conditions associated with the free values of the curve on the boundaries of the type space. The third condition is the usual complementary slackness condition associated with the optimal choice of the multiplier.

The last condition is technical. It provides a sufficient condition for the integration by parts argument that proves the optimality of the candidate and is thus similar to the first corner condition of Weierstrass (see, e.g., Mesterton-Gibbons (2009, Chapter 6)). In the context of our problem this condition requires that

$$H_{\dot{v}(\theta)}^\lambda (\theta, v(\theta), \dot{v}(\theta)) = \theta - \dot{v}(\theta) \left(2\lambda(\theta) + \frac{\dot{v}(\theta)\xi}{\sqrt{2v(\theta) - [\dot{v}(\theta)]^2}} \right), \tag{C.6}$$

be continuous throughout the type space, and, since the functions v and \dot{v} are both continuous, this requirement is equivalent to the property that $\dot{v}(\theta) = 0$ at every point of discontinuity of the Lagrange multiplier.

Proof of Lemma C.2. Assume that the conditions of the statement hold true and pick an arbitrary $w \in \mathcal{V}_1$. Combining Lemma C.4 below with Equation (C.5), the definition of \mathcal{V}_1 and the nonnegativity of the Lagrange multiplier we deduce that

$$\begin{aligned} \delta(v; w) &:= \int_0^{\theta_H} \left(F(\theta, w(\theta), \dot{w}(\theta)) - F(\theta, v(\theta), \dot{v}(\theta)) \right) d\theta \\ &\leq \int_0^{\theta_H} \left(H^\lambda(\theta, w(\theta), \dot{w}(\theta)) - H^\lambda(\theta, v(\theta), \dot{v}(\theta)) \right) d\theta \\ &\leq \int_0^{\theta_H} \left((w(\theta) - v(\theta))H_{v(\theta)}^\lambda(\theta) + (\dot{w}(\theta) - \dot{v}(\theta))H_{\dot{v}(\theta)}^\lambda(\theta) \right) d\theta \\ &= \sum_{n=0}^N \int_{\theta_n}^{\theta_{n+1}} \left((w(\theta) - v(\theta))H_{v(\theta)}^\lambda(\theta) + (\dot{w}(\theta) - \dot{v}(\theta))H_{\dot{v}(\theta)}^\lambda(\theta) \right) d\theta, \end{aligned} \tag{C.7}$$

where $\theta_0 = 0, \theta_{N+1} = \theta_H$, the N -tuple $(\theta_n)_{n=1}^N \in \text{int}(\Theta)^N$ identifies the points of discontinuity of the multiplier, and we have set

$$H_k^\lambda(\theta) = H_k^\lambda(\theta, v(\theta), \dot{v}(\theta)), \quad k \in \{v(\theta), \dot{v}(\theta)\}. \tag{C.8}$$

By assumption, we have that the functions v, \dot{v}, w , and λ are absolutely continuous on the interval (x_n, x_{n+1}) . Therefore, the functions $\dot{w} - \dot{v}$ and

$$H_{\dot{v}(\theta)}^\lambda(\theta) = F_{\dot{v}(\theta)}(\theta, v(\theta), \dot{v}(\theta)) - 2\dot{v}(\theta)\lambda(\theta), \tag{C.9}$$

are, respectively, Lebesgue integrable and absolutely continuous on Θ ; and we can thus integrate by parts on the right hand side of Equation (C.7) to obtain that

$$\begin{aligned} \delta(v; w) &\leq \sum_{n=0}^N \left\{ (w(\theta) - v(\theta))H_{\dot{v}(\theta)}^\lambda(\theta) \Big|_{\theta_n}^{\theta_{n+1}} \right. \\ &\quad \left. + \int_{\theta_n}^{\theta_{n+1}} (w(\theta) - v(\theta)) \left[H_{v(\theta)}^\lambda(\theta) - \frac{d}{d\theta} H_{\dot{v}(\theta)}^\lambda(\theta) \right] d\theta \right\} \\ &= \sum_{n=0}^N (w(\theta) - v(\theta))H_{\dot{v}(\theta)}^\lambda(\theta) \Big|_{\theta_n}^{\theta_{n+1}} \\ &= (w(\theta) - v(\theta))H_{\dot{v}(\theta)}^\lambda(\theta) \Big|_0^{\theta_H} - \sum_{n=1}^N (w(\theta_n) - v(\theta_n))\Delta H_{\dot{v}(\theta)}^\lambda(\theta_n) \\ &= (w(\theta) - v(\theta))H_{\dot{v}(\theta)}^\lambda(\theta) \Big|_0^{\theta_H} = 0, \end{aligned} \tag{C.10}$$

where the first equality follows from Equation (C.3), the second follows by expanding and then collapsing the sum, and the last follows from Equation (C.4). □

Remark C.3. Note that for the function

$$H_{\dot{v}(\theta)}^\lambda(\theta) = 2\dot{v}(\theta)\lambda(\theta) + F_{\dot{v}(\theta)}(\theta), \tag{C.11}$$

to be continuous it is necessary and sufficient that the optimal indirect utility be such that $\dot{v}(\theta) = 0$ at every point of discontinuity of the multiplier.

Lemma C.4. *Let $(\theta, \lambda) \in \Theta \times AC_p^*(\Theta; \mathbb{R}_+)$ be fixed. Then the map $(v, \dot{v}) \mapsto H^\lambda(\theta, v, \dot{v})$ is strictly concave on the set of pairs such that $c(v, \dot{v}) \geq 0$.*

Proof. This follows by verifying that the determinant and trace of the Hessian matrix are, respectively, strictly positive and strictly negative. We omit the details. □

Proof of Proposition 5.4. when $\gamma_1 \leq \xi/3$. To establish the result, it suffices to show that one can construct a Lagrange multiplier $\lambda \in AC_p^*(\Theta; \mathbb{R}_+)$ such that the pair (λ, v_1^*) satisfies the conditions

of Lemma C.2. As is easily seen, we have that the candidate optimizer belongs to $C_p^2(\Theta; \mathbb{R})$ and satisfies

$$c(v_1^*(\theta), v_1^*(\theta)) = 0, \quad \theta \in \Theta. \quad (\text{C.12})$$

Therefore, Equation (C.5) holds. On the other hand, substituting the candidate optimizer into Equations (C.3) and (C.4) shows that the Lagrange multiplier must be chosen in such a way that

$$\begin{aligned} 0 &= \frac{\gamma_1}{\xi - \gamma_1} + 2\lambda(\theta) - 2, & \theta &\leq \frac{\theta_H}{3} \\ 0 &= \frac{\gamma_1}{\xi - \gamma_1} + 2\lambda(\theta) - \frac{1}{2} + \lambda(\theta) \left(\theta - \frac{\theta_H}{3} \right), & \theta &> \frac{\theta_H}{3}, \end{aligned} \quad (\text{C.13})$$

$$(\text{C.14})$$

and

$$0 = \frac{\gamma_1}{\xi - \gamma_1} + 2\lambda(\theta_H) - \frac{1}{2}. \quad (\text{C.15})$$

A direct calculation shows that the unique solution to these equations is piecewise constant and explicitly given by

$$\lambda(\theta) := 1 - \frac{\gamma_1}{2(\xi - \gamma_1)} - \mathbf{1}_{\{3\theta > \theta_H\}} \frac{3}{4}. \quad (\text{C.16})$$

Since $\gamma_1 \leq \xi/3$, we have that $\lambda(\theta)$ is nonnegative for all $\theta \in \Theta$. Therefore, it now only remains to establish that the function

$$\theta \mapsto H_{v_1^*(\theta)}^\lambda(\theta, v_1^*(\theta), v_1^*(\theta)) = \theta - v_1^*(\theta) \left(2\lambda(\theta) + \frac{\xi}{\sqrt{2v_1^*(\theta) - [v_1^*(\theta)]^2}} \right), \quad (\text{C.17})$$

is continuous on Θ but this property follows from Remark C.3, the smoothness of the candidate optimizer, and the fact that $v_1^*(\theta_H/3) = 0$. \square

Lemma C.5. *For every $\gamma_1 \leq \gamma_1^*$, there exists a unique solution $(w, \theta^*) \in C_p^2(\Theta; \mathbb{R}) \times \Theta$ to the free boundary problem defined by Equations (48)–(50).*

Proof. We start by observing that (w, θ^*) is a solution to the free boundary problem if and only if $m(x) := w(x + \theta^*)$ solves the initial value problem

$$m(x)(1 + \dot{m}(x)) - [\dot{m}(x)]^2 = \frac{3}{2\xi} (2m(x) - [\dot{m}(x)]^2)^{\frac{3}{2}}, \quad (\text{C.18})$$

$$\dot{m}(0) = m(0) - \frac{1}{2}(\xi - \gamma_1)^2 = 0, \quad (\text{C.19})$$

and the constant $\theta^* \in \Theta$ solves

$$Q(\theta^*) := \theta_H - \frac{\xi \dot{m}(\theta_H - \theta^*)}{\sqrt{2m(\theta_H - \theta^*) - [\dot{m}(\theta_H - \theta^*)]^2}}. \tag{C.20}$$

Since $\gamma_1 \leq \gamma_1^*$, we have that

$$\frac{1}{2}(\xi - \gamma_1)^2 \geq \frac{1}{2}(\xi - \gamma_1^*)^2 = v^*(0) > \frac{1}{18}\xi^2. \tag{C.21}$$

Therefore, it follows from Lemma B.5 that the unique classical solution to Equation (C.18) subject to Equation (C.19) is given by $m(x) = v(x; q_1)$ with $q_1 = \frac{1}{2}(\xi - \gamma_1)^2$ and it now only remains to show that Equation (C.20) admits a unique solution.

Since the function Q is continuous on Θ and $Q(\theta_H) = \theta_H > 0$, the existence claim will follow from the intermediate value theorem once we show that $Q(0) \leq 0$. To this end, consider the continuously differentiable function

$$S(\theta) := \theta - \frac{\xi \dot{v}(\theta; q_1)}{\sqrt{2v(\theta, q_1) - [\dot{v}(\theta; q_1)]^2}}, \tag{C.22}$$

and observe that, since $Q(0) = S(\theta_H)$, it is sufficient to prove that $S(\theta) \leq 0$ in a left neighborhood of θ_H . Differentiating the above expression and using the fact that the function $v(\theta; q_1)$ solves (C.18) shows that

$$\dot{S}(\theta) = \frac{\xi}{\sqrt{2v(\theta, q_1) - [\dot{v}(\theta; q_1)]^2}} - 2. \tag{C.23}$$

To proceed further, we distinguish three cases. If the fee rate on the outside fund is such that $q_1 \in C_2$, then we know from the proof of Lemma B.5 that

$$\inf_{\theta \in \Theta} \sqrt{2v(\theta, q_1) - [\dot{v}(\theta; q_1)]^2} \geq \frac{2}{3}\xi. \tag{C.24}$$

This implies that we have $\dot{S}(\theta) \leq -\frac{1}{2}$ for all $\theta \in \Theta$ and the desired result follows by noting that $S(0) = 0$. Assume next that the fee rate is such that $q_1 \in [\frac{1}{8}, \frac{2}{9}]\xi^2$. In this case we know from Lemma B.5 that $\ddot{v}(\theta; q_1) \leq 1$ for all $\theta \in \Theta$. This implies that the derivative in Equation (C.23) is decreasing and the desired result follows by noting that

$$q_1 \geq \frac{1}{8}\xi^2 \implies \dot{S}(\theta) \leq \dot{S}(0) = \frac{\xi}{\sqrt{2q_1}} - 2 \leq 0 = S(0). \tag{C.25}$$

Finally, assume that the fee rate on the familiar asset is such that $q_1 < \frac{1}{8}\xi^2$, and denote by $q(\theta)$ the function that describes the unique strictly positive solution to Equation (B.74) in C_1 . As shown in the proof of Lemma B.6, this function is continuous, nonincreasing, and equal to $\frac{1}{8}\xi^2$ at the origin.

By continuity, this implies that we have $q_1 = q(\hat{\theta})$ for some strictly positive type $\hat{\theta}$ such that

$$S(\hat{\theta}) = \hat{\theta} - \frac{\xi \dot{v}(\hat{\theta}; q(\hat{\theta}))}{\sqrt{2v(\hat{\theta}, q(\hat{\theta})) - [\dot{v}(\hat{\theta}; q(\hat{\theta}))]^2}} = 0. \quad (\text{C.26})$$

On the other hand, since $q_1 < \frac{1}{8}\xi^2 < \frac{2}{9}\xi^2$ we know from Lemma B.5 that $\dot{v}(\theta; q_1) \leq 1$ for all $\theta \in \Theta$ and it follows that the derivative in Equation (C.23) is decreasing. Using this property in conjunction with the fact that

$$q_1 < \frac{1}{8}\xi^2 \implies \dot{S}(0) = -2 + \frac{\xi}{\sqrt{2q_1}} > 0 = S(\hat{\theta}), \quad (\text{C.27})$$

we then deduce that $\dot{S}(\hat{\theta}) < 0$, and it follows that $S(\theta) \leq S(\hat{\theta}) = 0$ for all $\theta \geq \hat{\theta}$. To complete the proof, it now remains to establish uniqueness. A direct calculation using Equation (C.20) and the fact that the function $v(\theta; q_1)$ solves Equation (C.18) implies that

$$Q'(\theta) = 3 - \frac{\xi}{\sqrt{2v(\theta, q_1) - [\dot{v}(\theta; q_1)]^2}} = 1 - \dot{S}(\theta). \quad (\text{C.28})$$

If $q_1 \geq \frac{1}{8}\xi^2$, then the first part of the proof shows that we have $\dot{S}(\theta) \leq 0$ for all $\theta \in \Theta$. This implies that the function $Q(\theta)$ is strictly increasing and the required uniqueness follows. On the other hand, if we have $v^*(0) < q_1 < \frac{1}{8}\xi^2$, then we know from the first part of the proof that $Q'(\theta)$ is decreasing and has initial value

$$Q'(0) = 3 - \frac{\xi}{\sqrt{2q_1}} \geq 3 - \frac{\xi}{\sqrt{2\xi^2/18}} = 0. \quad (\text{C.29})$$

Two cases may then occur: either $Q'(\theta_H) \geq 0$, in which case, the function $Q(\theta)$ is strictly increasing, or $Q'(\theta_H) < 0$, in which case the function $Q(\theta)$ is inverse U -shaped with a maximum whose value exceeds $Q(\theta_H)$. The required uniqueness follows by observing that in both cases, the function crosses the horizontal axis only once. \square

Proof of Proposition 5.4. when $\gamma_1 \in (\xi/3, \gamma_1^*]$. To establish the result, it suffices to show that one can construct a Lagrange multiplier $\lambda \in AC_p^*(\Theta; \mathbb{R}_+)$ such that the pair (λ, v_1^*) satisfies the conditions of Lemma C.2. As is easily seen, we have that the candidate optimizer belongs to $C_p^2(\Theta; \mathbb{R})$ and satisfies

$$\begin{aligned} c(v_1^*(\theta), \dot{v}_1^*(\theta)) &= \mathbf{1}_{\{\theta > \theta^*\}} (2w(\theta) - [\dot{w}(\theta)]^2 - 2q_1) \\ &= \mathbf{1}_{\{\theta > \theta^*\}} (2v(\theta - \theta^*; q_1) - [\dot{v}(\theta - \theta^*; q_1)]^2 - 2q_1), \end{aligned} \quad (\text{C.30})$$

where the second equality follows from the construction in the proof of Lemma C.5. Since $\gamma_1 \geq \frac{1}{3}\xi$, we have $q_1 \leq \frac{2}{9}\xi^2$. Therefore, the result of Lemma B.5 implies that the right hand side of the

previous expression is strictly increasing in $\theta \in [\theta^*, \theta_H]$ and it follows that we have

$$c(v_1^*(\theta), v_1^*(\theta)) > c(v_1^*(\theta^*), v_1^*(\theta^*)) = 2(v(0; q_1) - q_1) = 0, \quad \theta > \theta^*, \quad (\text{C.31})$$

which establishes that the candidate optimizer lies in \mathcal{V}_1 and shows that a necessary condition for Equation (C.5) is that $\lambda(\theta) = 0$ for all $\theta > \theta^*$. On the other hand, using the fact that the function $v(\theta; q)$ solves Equation (48) and substituting into Equations (C.3) and (C.4) shows that the Lagrange multiplier must be chosen in such a way that

$$0 = \frac{\gamma_1}{\xi - \gamma_1} + 2\lambda(\theta) - 2, \quad \theta \leq \theta^*. \quad (\text{C.32})$$

Solving that equation shows the Lagrange multiplier is given by

$$\lambda(\theta) := \mathbf{1}_{\{\theta \leq \theta^*\}} \left(1 - \frac{\gamma_1}{2(\xi - \gamma_1)} \right). \quad (\text{C.33})$$

Since $\gamma_1 \leq \xi/3$, we have that $\lambda(\theta)$ is nonnegative for all $\theta \in \Theta$. Therefore, it now only remains to establish that the function

$$\theta \mapsto H_{v_1^*(\theta)}^\lambda(\theta, v_1^*(\theta), v_1^*(\theta)) = \theta - v_1^*(\theta) \left(2\lambda(\theta) + \frac{\xi}{\sqrt{2v_1^*(\theta) - [v_1^*(\theta)]^2}} \right), \quad (\text{C.34})$$

is continuous on Θ but this property follows from Remark C.3, the smoothness of the candidate optimizer, and the fact that $v_1^*(\theta^*) = v(0; q_1) = 0$. \square

C.3 | The optimal fund menu

Proof of Theorem 5.5. Since $v_1^* \in C_p^2(\Theta; \mathbb{R})$, we have $\phi_1^* \in AC(\Theta; \mathbb{R}^2)$. Therefore, Lemma 5.3 and Proposition 5.4 imply that

$$M_1 = \sup_{\phi \in \Phi} I(\phi) \leq \frac{V_1}{\theta_H} = \int_{\Theta} F(\theta, v_1^*(\theta), v_1^*(\theta)) \frac{d\theta}{\theta_H} = I(\phi_1^*), \quad (\text{C.35})$$

and the statement will follow once we show that the fund loading function ϕ_1^* belongs to Φ_1 and satisfies Equation (52). We distinguish two cases depending on the fee rate of the outside fund. Assume first that we have $\gamma_1 \leq \frac{1}{3}\xi$ so that the function $v_1^*(\theta)$ is given by Equation (47). In this case, the result follows by observing that we have

$$F(\theta, v_1^*(\theta), v_1^*(\theta)) = \frac{\gamma_1}{a}(\xi - \gamma_1) + \frac{\theta^*}{a}(\theta - \theta^*)_+ > 0, \quad (\text{C.36})$$

as well as

$$\xi - \gamma_1 = \frac{\phi_{11}^*(\theta)}{\|\phi_1^*(\theta)\|^2} (\phi_1^*(\theta)^\top \xi(\theta) - 1)_+, \quad (\text{C.37})$$

and

$$\begin{aligned} & \phi_1^*(\theta')^\top \xi(\theta) - 1 - \frac{\phi_1^*(\theta)^\top \phi_1^*(\theta')}{\|\phi_1^*(\theta)\|^2} (\phi_1^*(\theta)^\top \xi(\theta) - 1)_+ \\ &= \mathbf{1}_{\{\theta \leq \theta^* < \theta'\}} \frac{(\theta^* - \theta)(\theta^* - \theta')}{\gamma_1(\xi - \gamma_1) + \theta^*(\theta' - \theta^*)} \leq 0 \end{aligned} \quad (\text{C.38})$$

for all $(\theta, \theta') \in \Theta^2$. Assume next that the fee rate $\gamma_1 \in (\frac{1}{3}\xi^2, \gamma_1^*]$, and let us start by establishing the validity of Equation (52). Since $v_1^*(\theta)$ is piecewise smooth, we have that the mapping defined by

$$\theta \mapsto F_1^*(\theta) := F(\theta, v_1^*(\theta), \dot{v}_1^*(\theta)), \quad (\text{C.39})$$

is absolutely continuous. On the other hand, since $\gamma_1 > \frac{1}{3}\xi$, it follows from Equation (51) and the proofs of Lemmas B.5 and C.5 that we have

$$\min\{\dot{v}_1^*(\theta), 1 - \dot{v}_1^*(\theta), \dot{v}_1^*(\theta), \dot{v}_1^*(\theta)\} \geq 0, \quad \theta \in [\theta^*, \theta_H]. \quad (\text{C.40})$$

Using this property in conjunction with the fundamental theorem of calculus, we then deduce that

$$\begin{aligned} \dot{F}_1^*(\theta) &= \mathbf{1}_{\{\theta \geq \theta^*\}} \left(\theta \dot{v}_1^*(\theta) - \dot{v}_1^*(\theta) + \dot{v}_1^*(\theta)(1 - \dot{v}_1^*(\theta)) \frac{\xi}{g_1^*(\theta)} \right) \\ &\geq \mathbf{1}_{\{\theta \geq \theta^*\}} (\theta \dot{v}_1^*(\theta) - \dot{v}_1^*(\theta)) = \int_{\theta^*}^{\max\{\theta, \theta^*\}} (\dot{v}_1^*(\theta) - \dot{v}_1^*(x)) dx \geq 0. \end{aligned} \quad (\text{C.41})$$

for the function

$$g_1^*(\theta) := \sqrt{2v_1^*(\theta) - [v_1^*(\theta)]^2}. \quad (\text{C.42})$$

This shows that the absolutely continuous function $F_1^*(\theta)$ is nondecreasing throughout the type space and Equation (52) now follows by observing that

$$F(\theta, v_1^*(\theta), \dot{v}_1^*(\theta)) = \gamma_1(\xi - \gamma_1) > 0, \quad \theta \in [0, \theta^*]. \quad (\text{C.43})$$

To complete the proof, we now need to show that the fund loading function $\phi_1^*(\theta)$ is incentive compatible. A direct calculation using Equation (53) shows that

$$\xi - \gamma_1 - \frac{\phi_{11}^*(\theta)}{\|\phi_1^*(\theta)\|^2} (\phi_1^*(\theta)^\top \xi(\theta) - 1)_+ = \mathbf{1}_{\{\theta \geq \theta^*\}} (g_1^*(\theta^*) - g_1^*(\theta)), \quad (\text{C.44})$$

and Equation (45) follows by noting that, as a result of Equation (C.40), the function $g_1^*(\theta)$ is nondecreasing on the interval $[\theta^*, \theta_H]$. On the other hand, using Equation (C.40) and proceeding

as in the proof of Theorem 3.5 shows that the validity of Equation (10) is equivalent to

$$g_1^*(\theta) \geq h(\theta, \theta') := \frac{2v_1^*(\theta') + (\theta - \theta' - v_1^*(\theta))v_1^*(\theta')}{g_1^*(\theta')}, \quad (\theta, \theta') \in \Theta^2. \quad (\text{C.45})$$

To prove this inequality, we start by decomposing the set Θ^2 into the union of the disjoint subsets $(\Theta_i)_{i=1}^4$ defined by

$$\begin{aligned} \Theta_1 &:= \{(\theta, \theta') \in \Theta^2 : \max\{\theta, \theta'\} \leq \theta^*\}, \\ \Theta_2 &:= \{(\theta, \theta') \in \Theta^2 : \theta \leq \theta^* \text{ and } \theta' > \theta^*\}, \\ \Theta_3 &:= \{(\theta, \theta') \in \Theta^2 : \theta > \theta^* \text{ and } \theta' \leq \theta^*\}, \end{aligned} \quad (\text{C.46})$$

and

$$\Theta_4 := \{(\theta, \theta') \in \Theta^2 : \min\{\theta, \theta'\} > \theta^*\}. \quad (\text{C.47})$$

On the set Θ_1 , the inequality holds since

$$g_1^*(\theta) = h(\theta, \theta') = \sqrt{2q_1} = \xi - \gamma_1, \quad (\theta, \theta') \in \Theta_1. \quad (\text{C.48})$$

On the set Θ_3 , the inequality boils down to

$$g_1^*(\theta) \geq g_1^*(\theta^*) = \sqrt{2q_1} = \xi - g_1, \quad \theta > \theta^*, \quad (\text{C.49})$$

which is satisfied because $g_1^*(\theta)$ is nondecreasing on $[\theta^*, \theta_H]$ as a result of Equation (C.40). On the set Θ_2 , we have that

$$g_1^*(\theta) - h(\theta, \theta') = \xi - \gamma_1 - \frac{(\xi - \gamma_1)^2 + (\theta - \theta')v_1^*(\theta')}{g_1^*(\theta')}, \quad (\text{C.50})$$

is strictly decreasing with respect to θ and it follows that the validity of Equation (C.45) on that set is equivalent to

$$h(\theta^*, \theta') \leq \xi - \gamma_1, \quad \theta' > \theta^*. \quad (\text{C.51})$$

Differentiating the right hand side of Equation (C.45) gives

$$\frac{\partial h}{\partial \theta'}(\theta, \theta') = (\theta - v_1^*(\theta) - \theta' + v_1^*(\theta')) \left[\frac{2v_1^*(\theta')v_1^*(\theta') - [v_1^*(\theta')]^2}{g_1^*(\theta')^3} \right]. \quad (\text{C.52})$$

Combining Equation (C.40) with the fundamental theorem of calculus, we deduce that

$$2v_1^*(\theta')v_1^*(\theta') - [v_1^*(\theta')]^2 = \int_{\theta^*}^{\theta'} 2v_1^*(x)v_1^*(x)dx \geq 0, \quad (\text{C.53})$$

for all $\theta' \geq \theta^*$. On the other hand, since $v_1^*(\theta^*) = 0$, it follows from Equation (C.40) and the fundamental theorem of calculus that

$$\theta^* - \theta' + v_1^*(\theta') = \theta^* - \theta' + \int_{\theta^*}^{\theta'} v_1^*(x) dx \leq 0, \tag{C.54}$$

for all $\theta' \geq \theta$. This shows that $h(\theta^*, \theta')$ is decreasing in θ' and Equation (C.51) now follows by observing that

$$h(\theta^*, \theta^*) = \frac{2v_1^*(\theta^*) - [v_1^*(\theta^*)]^2}{g_1^*(\theta^*)} = g_1^*(\theta^*) = \xi - \gamma_1. \tag{C.55}$$

Consider finally the set Θ_4 . Since $h(\theta, \theta) = g_1^*(\theta)$, it is sufficient to show that for any fixed $\theta > \theta^*$, the function $h(\theta, \theta')$ reaches a maximum over $(\theta^*, \theta_H]$ at the point $\theta' = \theta$. In view of Equation (C.53), we have that the sign of $\frac{\partial h}{\partial \theta'}$ is determined by the sign of

$$(\theta - v_1^*(\theta)) - (\theta' - v_1^*(\theta')). \tag{C.56}$$

By Equation (C.40), we have that $x - v_1^*(x)$ is nondecreasing on $[\theta^*, \theta_H]$ and it follows that the above expression is nonnegative if and only if $\theta' \leq \theta$. □

Proof of Proposition 5.6. Arguments similar to those of Sections 3.2 and 5.2 show that under exclusivity the value function of the manager satisfies

$$\theta_H M_1 \leq \theta_H M_{1,E} \leq V_{1,E} \equiv \sup_{v \in \mathcal{V}_{1,E}} \int_{\Theta} F(\theta, v(\theta), v(\theta)) d\theta, \tag{R_{1,E}}$$

where $\mathcal{V}_{1,E}$ denotes the set of functions $v \in AC(\Theta; \mathbb{R})$ that satisfy Equation (54). Consider the candidate optimizer $v_{1,E}^*(\theta)$ defined in the statement. As is easily seen, we have that this function is absolutely continuous. On the other hand, the construction of the function $w(\theta)$ implies that

$$\mathbf{1}_{\{\theta \geq \theta^*\}}(w(\theta) - v(\theta - \theta^*; q_1)) = 0 \quad \text{with } q_1 \equiv \frac{1}{2}(\xi - \gamma_1)^2. \tag{C.57}$$

Therefore, it follows from Lemma B.5 that $v_{1,E}^*(\theta)$ is nondecreasing on $[\theta^*, \theta_H]$ and, since $v_{1,E}^*(\theta) = q_1$ for all $\theta \leq \theta^*$, we conclude that $v_{1,E}^* \in \mathcal{V}_{1,E}$. To show that it attains the supremum consider the multiplier

$$\lambda_E(\theta) \equiv \mathbf{1}_{\{\theta \leq \theta^*\}} \left(3 - \frac{\xi}{\xi - \gamma_1} \right) \geq 0, \tag{C.58}$$

where the inequality follows from the fact that $\gamma_1 \leq \gamma_1^* \leq \frac{2}{3}\xi$. A direct calculation shows that the pair $(v_{1,E}^*, \lambda_E)$ satisfies all the conditions of Lemma C.6 below and it thus follows that we have

$$V_{1,E} = \int_{\Theta} F(\theta, v_{1,E}^*(\theta), v_{1,E}^*(\theta)) d\theta. \tag{C.59}$$

To complete the proof, we distinguish two cases depending on the level of the fee rate on the familiar asset. Assume first that $\gamma_1 > \frac{1}{3}\xi$. In this case

$$0 = \left| v_1^*(\theta) - v_{1,E}^*(\theta) \right| = \left\| \phi_1^*(\theta) - \phi_{1,E}^*(\theta) \right\|, \quad \theta \in \Theta, \tag{C.60}$$

and the desired result now follows from $(\mathcal{R}_{1,E})$ and Theorem 5.5. Assume next that the fee rate $\gamma \leq \frac{1}{3}\xi$ and consider the fund loading function defined in the statement. To complete the proof, we now show that this function belongs to $\Phi_{1,E}$ but not to Φ_1 . To establish the former, we need to show that

$$\max \left\{ q_1, \sup_{\theta' \in \Theta} \frac{1}{2} \left(\frac{\phi_{1,E}^*(\theta')^\top \xi(\theta) - 1}{\|\phi_{1,E}^*(\theta')\|} \right)_+^2 \right\} \leq v_{1,E}^*(\theta), \quad \theta \in \Theta. \tag{C.61}$$

but, since $v_{1,E}^*(\theta) \geq q_1$ for all $\theta \in \Theta$, we have that the validity of this inequality is equivalent to the requirement that

$$B(\theta, \theta') \equiv 4v_{1,E}^*(\theta)v_{1,E}^*(\theta') - \left(2v_{1,E}^*(\theta') + (\theta - \theta')v_{1,E}^*(\theta') \right)_+^2, \tag{C.62}$$

be nonnegative for all $(\theta, \theta') \in \Theta \times (\theta^*, \theta_H]$. Differentiating this function with respect to its first argument gives

$$\frac{dB}{d\theta} = \mathbf{1}_{\{\theta > \theta^*\}} 4\dot{v}_{1,E}^*(\theta)v_{1,E}^*(\theta') - 2\dot{v}_{1,E}^*(\theta') \left(2v_{1,E}^*(\theta') + (\theta - \theta')v_{1,E}^*(\theta') \right)_+. \tag{C.63}$$

Since the fee rate $\gamma_1 < \frac{1}{3}\xi$, we have that $q_1 \in C_2$. Therefore, it follows from Lemma B.5 that for all types $\theta' > \theta^*$ we have

$$\ddot{v}_{1,E}^*(\theta') = \ddot{v}(\theta' - \theta^*, q_1) \geq 1, \tag{C.64}$$

and

$$\dot{v}_{1,E}^*(\theta') = \dot{v}_{1,E}^*(\theta^*) + \int_{\theta^*}^{\theta'} \ddot{v}_{1,E}^*(\theta) d\theta \geq \theta' - \theta^*. \tag{C.65}$$

Combining these properties with Equation (50) shows that for all $(\theta, \theta') \in [\theta^*, \theta_H]^2$ we have

$$\begin{aligned} 2v_{1,E}^*(\theta') + (\theta - \theta')\dot{v}_{1,E}^*(\theta') &\geq 2v_{1,E}^*(\theta') + (\theta^* - \theta')\dot{v}_{1,E}^*(\theta') \\ &\geq 2v_{1,E}^*(\theta') - [\dot{v}_{1,E}^*(\theta')]^2 \\ &\geq 2v_{1,E}^*(\theta_H) - [\dot{v}_{1,E}^*(\theta_H)]^2 = \left(\frac{\xi}{\theta_H} \right)^2 [\dot{v}_{1,E}^*(\theta_H)]^2 \geq 0, \end{aligned} \tag{C.66}$$

and therefore

$$\begin{aligned}
 (\theta - \theta') \frac{dB}{d\theta} &= 2(\theta - \theta')^2 \left(2v_{1,E}^*(\theta') \frac{\dot{v}_{1,E}^*(\theta) - \dot{v}_{1,E}^*(\theta')}{\theta - \theta'} - [v_{1,E}^*(\theta')]^2 \right) \\
 &= 2(\theta - \theta')^2 \left(2v_{1,E}^*(\theta') - [\dot{v}_{1,E}^*(\theta')]^2 \right) \geq 0.
 \end{aligned}
 \tag{C.67}$$

This shows that for any given $\theta' > \theta^*$ the function $\theta \mapsto B(\theta, \theta')$ reaches the minimum of zero over $[\theta^*, \theta_H]$, and we now have to consider types such that $\theta \leq \theta^* < \theta'$. For such types we have that $\frac{dB}{d\theta}$ is negative or zero, and the desired property now follows from the fact that, as shown above, we have $B(\theta^*, \theta') \geq 0$ for all $\theta' > \theta^*$.

To complete the proof, it now remains to show that we have $\phi_{1,E}^* \notin \Phi_{1,E}$. Proceeding as in the proof of Theorem 5.5, we have that on $[\theta^*, \theta_H]^2$ the validity of the nonexclusive incentive compatibility condition is equivalent to

$$g_{1,E}^*(\theta) \geq h(\theta, \theta') \equiv \frac{2v_{1,E}^*(\theta') + (\theta - \theta' - \dot{v}_{1,E}^*(\theta))\dot{v}_{1,E}^*(\theta')}{g_{1,E}^*(\theta)},
 \tag{C.68}$$

with

$$g_{1,E}^*(\theta) \equiv \sqrt{2v_{1,E}^*(\theta) - [\dot{v}_{1,E}^*(\theta)]^2}.
 \tag{C.69}$$

Combining Equations (C.64) and (C.65), we deduce that

$$2v_{1,E}^*(\theta)\dot{v}_{1,E}^*(\theta) - [\dot{v}_{1,E}^*(\theta)]^2 \geq 2v_{1,E}^*(\theta) - [\dot{v}_{1,E}^*(\theta)]^2 \geq 0, \quad \theta \geq \theta^*.
 \tag{C.70}$$

Therefore, the sign of

$$\frac{dh}{d\theta'} = \left(\theta - \dot{v}_{1,E}^*(\theta) - \theta' + \dot{v}_{1,E}^*(\theta') \right) \left(\frac{2v_{1,E}^*(\theta')\dot{v}_{1,E}^*(\theta') - [\dot{v}_{1,E}^*(\theta')]^2}{g_{1,E}^*(\theta')^3} \right),
 \tag{C.71}$$

is determined by the sign of the first bracket on the right. Because of Equation (C.65), we have that the function $\theta - \dot{v}_{1,E}^*(\theta)$ is decreasing. This implies that

$$(\theta - \theta') \frac{dh}{d\theta'}(\theta, \theta') \leq 0, \quad \theta^* \leq \min\{\theta, \theta'\},
 \tag{C.72}$$

and using this inequality in conjunction with the fact that $h(\theta, \theta) = g_{1,E}^*(\theta)$, we deduce that $g_{1,E}^*(\theta) < h(\theta, \theta')$ for all $\theta \neq \theta'$ in $[\theta^*, \theta_H]^2$. This shows that the inequality in Equation (C.68) fails and the desired result follows. \square

Consider the function defined by

$$G^\lambda(\theta, v, p) \equiv F(\theta, v, p) + \lambda(\theta)(v - q_1),
 \tag{C.73}$$

and observe that, as a result of Lemma B.4, this function is strictly concave in v and p . The following lemma is the counterpart of Lemma C.2 for the case where the manager can force investors to commit to a single fund.

Lemma C.6. *Let $(v, \lambda) \in \mathcal{V}_{1,E} \times AC_p^*(\Theta; \mathbb{R}_+)$ be such that $\dot{v} \in AC(\Theta; \mathbb{R})$, and denote by \mathcal{C} the set of points where the function λ is continuous. If*

$$\left(G_{v(\theta)}^\lambda - \frac{d}{d\theta} G_{\dot{v}(\theta)}^\lambda \right) (\theta, v(\theta), \dot{v}(\theta)) = 0, \quad \theta \in \mathcal{C}, \quad (\text{C.74})$$

$$G_{\dot{v}(\theta)}^\lambda (\theta, v(\theta), \dot{v}(\theta)) = 0, \quad \theta \in \{0, \theta_H\}, \quad (\text{C.75})$$

$$\lambda(\theta)(v(\theta) - q_1) = 0, \quad \theta \in \Theta, \quad (\text{C.76})$$

and $G_{\dot{v}(\theta)}^\lambda (\theta, v(\theta), \dot{v}(\theta))$ is continuous, then v attains the supremum in $(\mathcal{R}_{1,E})$.

Proof. The proof is similar to that of Lemma C.2. We omit the details. □