Default penalties and fee structure in private equity partnerships[1]

Albert Banal-Estañol
Universitat Pompeu Fabra & Barcelona GSE
albert.banalestanol@upf.edu

Filippo Ippolito*
Universitat Pompeu Fabra & Barcelona GSE
filippo.ippolito@upf.edu

Sergio Vicente
Universidad Carlos III de Madrid
sergio.vicente@uc3m.es

May 8, 2018

Abstract

Default penalties are commonly observed in private equity funds. These penalties are levied on limited partners that miss out on a capital call. We construct a two-period model of optimal contracting between a limited partner and general partners whose managerial ability and diligence in running the fund are unobservable. We show that default penalties are used in combination with fees deferrals, insulating general partners against partnership interruptions, to reduce the distortions in investment scales caused by information asymmetries between the limited and the general partners. We also show that the optimal fee structure entails management fees that are proportional to the fund under management, performance fees that are paid when the funds deliver high returns, and an extra set of fees that reward managerial ability.

JEL Classifications: D82, D86, G23, G24, J33
Keywords: private equity, default penalties, management fees, transaction fees

[1]We are especially grateful to Andrés Almazán, Ulf Axelson, Catherine Casamatta, Xavier Freixas, Javier Gil-Bazo, Alexander Guembel, Demian Macedo, David Martinez-Miera, Alan Morrison, Ludovic Phalippou, Berto Portichol, Pablo Ruiz-Verdú (discussant), Armin Schwienbacher, Per Strömberg, and Anna Toldrà for useful comments and feedback. We also thank seminar participants at Instituto de Empresa, Pompeu Fabra, Cass Business School, the MOVE workshop on venture capital (Barcelona), the Corporate Finance Day in Ghent and the Finance Forum of the Spanish Finance Association. Sergio Vicente acknowledges support from Fundación Ramón Areces. All remaining errors are our own. A previous version of this paper was circulated under the title "Contractual Terms and Conditions in Private Equity Partnership Agreements."

*Contact author. Address: Filippo Ippolito, Department of Economics and Business and Barcelona Graduate School of Economics, Universitat Pompeu Fabra, Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain Tel. (+34) 93 542 2578, Fax. (+34) 93 542 1746

1 Introduction

Private equity partnerships are commonly structured as closed-end funds with a limited tradeability of shares, and a fixed life of about 10 years. The investors, known as *Limited Partners (LPs)*, commit capital at the fund's inception. The fund's managers, known as *General Partners (GPs)*, identify investment opportunities and request capital from LPs via *capital calls*. Since LPs supply the committed capital in stages, they retain the real option to default on their obligation. An LP may default on a capital call for several reasons, such as liquidity problems or as a means of reallocating its portfolio. Defaults on capital calls are not uncommon and can even be widespread.[2]

Because the installment practice is nearly universal, private equity partnership agreements typically specify a *default penalty* for the LP's failure to honor a capital call. Default penalties are often written as long lists of punishments, ranging from relatively mild to very severe, implying the loss of some or all of the profits, and the forfeiture of the defaulter's entire stake in the fund (Litvak (2004); Lerner, Hardymon, and Leamon (2005)).[3] In one of the most complete surveys of private equity partnership agreements to date, Toll and Vayner (2012) report that a high percentage of domestic (USA) venture (73%) and buyout funds (72%) include the most severe default penalty: forfeiture of a portion of the capital balance. This provision is less common but still prevalent in international funds (47%). As recognized by leading lawyers (e.g., Stone (2009)), even large LPs, such as pension funds or other institutional investors, agree to include significant default penalties in many of the small private equity funds (Litvak

[2]For anecdotal evidence see Stone (2009); Brett Byers, "Secondary Sales of Private Equity Interests," Venture Capital Fund of America (Feb 2002); R. Lindsay, I. Ashman and V. Hazelden, "Cayman Islands: Defaulting Limited Partners: Challenges for Private Equity in 2009," Walkers (Feb 2009); J. D. Corelli and S. Pindyck, "Capital Call Defaults Can Have Severe Consequences for Funds," Pepper Hamilton LLP (Apr 2010); S. G. Caplan, A. McWhirter and A. M. Ostrognai, "Private Equity Funds: Should You Be Thinking About Limited Partner Defaults?", Debevoise & Plimpton LLP (Feb 2009); Erin Griffith, "LP Defaults: What Exactly Happens?" The PE Hub Network (May 2009).

[3]"In the event that an investor defaults on its capital contribution obligations, private equity funds typically offer the sponsor broad flexibility in choosing from a laundry list of remedies. Delaware law (and the laws of many other jurisdictions) permits a fund to impose as stringent a remedy as the complete forfeiture of a defaulting partner's interests, if desired. Many funds provide for this possibility, while others choose a lesser but still potentially punitive forfeiture level." (Breslow and Schwartz (2015)).

(2004)). To our knowledge, and despite their pervasiveness and importance, the finance literature has not formally analyzed the role of default penalties in private equity partnerships, how harsh they should be, or on which fund characteristics they should depend on.

This paper offers a formal analysis of the role of default penalties in the design of private equity funds. We construct a two-period model with adverse selection and moral hazard in which an LP and a GP contract on the capital to invest in each period, the fees received by the GP, and the penalty borne by the LP if it misses a capital call. The model features three key ingredients. First, the fund is structured as a two-period relationship without commitment. The LP and the GP contract over two periods, but renegotiation and partnership interruptions are possible in the interim. Second, there is adverse selection and moral hazard in the relationship. On the one hand, there are better and worse GPs. On the other hand, the GP may run the fund diligently to generate high returns or adopt a passive strategy. The GP's type is private information, and the GP's actions are not observable. Third, an alternative investment opportunity may become available to the LP in the second period, giving rise to a potential partnership interruption in the interim.[4]

The optimal contract displays better GPs running funds with larger capital investments, collecting higher fees, and getting a higher compensation for partnership interruptions. Moreover, the optimal contract features fees that are proportional to the capital under management (*management fees*), fees that reward GPs for delivering high returns (*performance fees*), and an additional set of fees that reward high managerial ability and that are non-proportional to the fund capital (*extra fees*).[5] Default penalties constitute an integral part of the opti-

[4]When choosing to default on an obligation to a fund, an LP must also consider the effects on his reputation. For simplicity, we design the optimal contract abstracting from reputational considerations, and we assume that the financial penalty is the only cost that the LP bears. Nonetheless, several commentators argue that the stigma of failing to meet a capital call has diminished greatly since the financial crisis (Stone (2009)). Harris (2010) argues that LPs, which are frequently pension funds and other institutional investors–and not individuals–, are not necessarily in danger of losing access to a wide range of alternative private equity investment options after missing a capital call.

[5]In addition to management and performance fees, GPs are rewarded in reality with a variety of other fees, such as transaction, consulting, advisory and other related fees charged to portfolio companies. In our stylized model, we dub "extra fees" all payments that GPs perceive other than management and performance fees.

mal contract, improving the efficiency of investment decisions by insulating GPs against the possibility of a partnership interruption in the interim.

Most prior work on the contractual terms of private equity funds has focused on management and performance fees. The purpose of management fees in private equity funds is to cover the costs of managing the fund and typically comprise a fixed percentage of the committed capital.[6] We show that management fees should constitute a smaller percentage of capital under management for larger funds. This prediction is consistent with the finding of Gompers and Lerner (1999a) that management fees decrease with fund size. Performance-based fees, such as carried interest, are meant to reduce the moral hazard problems arising from fund mismanagement by the GP (Lerner and Schoar (2004)). In our framework, performance fees are used to incentivize the GP to run the fund diligently and, accordingly, reward the GP with a share of the fund returns when the fund performs well. Consistently with the findings in Toll and Vayner (2012), our model predicts that performance fees should be larger in environments in which the return of the fund is more uncertain and the tasks involved in managing the fund are more complex.

The role of the rest of fees in private equity partnerships, which are rarely fully rebated against management fees, is significantly less understood. As Metrick and Yasuda (2010a) put it, "it is not clear what these transaction fees are paying for since GPs should be already be receiving [...] management fees."[7] Phalippou, Rauch, and Umber (2018) and Legath (2011) highlight the importance of transaction fees in the LBO industry. They show that fees, other than management and performance fees, represent an important source of revenues for GPs and are computed as a non-trivial percentage of the size of the deals. In our framework,

[6] Management fees are generally regarded as a way to compensate the GP for the cost of running the fund. Monk and Sharma (2015), for example, state that: "It is widely perceived that management fees should just cover the cost of running the fund on a day-to-day basis, as opposed to providing a source of profit for the manager." In the comprehensive survey on PE/VC Partnership Agreements, Toll and Vayner (2012) report that "management fees go to pay the salaries and bonuses of the partners and supporting employees, to pay for office space, and to pay for other expenses related to operating the partnership".

[7] Transaction fees have recently attracted attention in the media because they have increased substantially in the aftermath of the crisis, while management and performance-based fees have been reduced. See The Economist, "Private equity: Fee high so dumb. Some buy-out firms' fees have gone up," November 12, 2011.

the role of extra fees in optimal contracts consists in rewarding better GPs for their higher managerial skills. To induce better GPs to run larger funds with lower proportional fees, LPs offer them larger non-proportional extra fees.

To our knowledge, our paper provides a first rationale for the use of default penalties in private equity partnerships. We show that default penalties constitute an instrument to reduce investment distortions in private equity funds. Contracts without default penalties must include high up-front payments to reward better GPs because the prospects of contract renegotiation reduce the fees that the GP can effectively perceive in the second period. But large payments in the first period would attract worse GPs pretending to be better GPs, as worse GPs could potentially collect high fees in the first period and refuse to run the fund in the second period. To prevent this type of behavior, LPs have to overinvest in funds run by better GPs and underinvest in funds run by worse GPs. Default penalties help reduce the need to distort investment scales because they constitute a vehicle to defer payments to the second period, instead of paying the extra fees up-front, and thus preclude worse GPs from running funds intended for better GPs.

In our framework, optimally-set default penalties amount to the entire amount of extra fees that better GPs obtain for their higher managerial ability. Consequently, default penalties are only included in contracts with better GPs and fully insulate them against partnership interruptions by compensating them for the entire amount of extra fees that they cease to collect when a partnership is broken. Consistent with the predictions of our model, the combined evidence of Litvak (2004) and Litvak (2009) shows that default penalties are higher in larger funds, that larger funds are run by better performing GPs, and that default penalties increase in the "option term," which is a measure of the relative importance of later capital calls versus earlier capital calls.

The literature on private equity has dedicated a great deal of attention to the relationship between venture capitalists and entrepreneurs (Gompers (1995), Hellmann (1998), Casamatta (2003), Cornelli and Yosha (2003), Kaplan and Strömberg (2003), Schmidt (2003), and Kaplan and Strömberg (2004)). However, there has been little research on the design of partnership agreements, with the exception of Axelson, Strömberg, and Weisbach (2009), which we refer to as ASW hereafter.[8] ASW show how committing capital for multiple investments reduces the GP's incentives to make bad investments. Relative to financing each deal separately, compensating a GP on aggregate returns reduces its incentives to invest in bad deals, since bad deals contaminate its stake in good deals. Instead, we focus on the problem of screening GPs with heterogeneous abilities and show how the screening problem is affected by the dynamic nature of the relationship between GPs and LPs.

In terms of the theory, we borrow ideas from dynamic models of adverse selection and moral hazard from Laffont and Tirole (1988), Rey and Salanie (1990), Laffont and Tirole (1990), Gibbons and Murphy (1992), Lambert (1993), Malcomson and Spinnewyn (1998), Dewatripont, Jewitt, and Tirole (1999), and Holmström (1999). The main novelty of our paper relies on the changing investment opportunities of the principal in the form of a random outside opportunity in the second period. In models without commitment, the prospects of renegotiation typically induce distortions of the sort found in this paper. We show that these distortions are reduced when partnerships can be interrupted because one of the parties find better investment alternatives. Moreover, we show that default penalties can reduce these distortions further by insulating the harmed party against partnership ruptures.

The rest of the paper is organized as follows. We describe the model in section 2. Section 3 lays out a description of benchmark one-shot contracts, which constitute the building blocks of the two-period contracts of our interest. In Section 4 we construct the optimal two-period contracts in a setup with zero default penalties for missing capital calls. In section 5 we characterize the optimal contracts when penalties can be included and address their role in reducing distortions. We conduct a comparison of separating and pooling contracts in Section 6. In Section 7 we provide a series of testable predictions generated by the model, relate them to current available evidence, and suggest potential empirical tests for further research.

[8]See Gompers and Lerner (1999b) and Sahlman (1990) for an overview of the structure and main characteristics of private equity partnerships.

Section 8 concludes. The Appendix contains a formal characterization of the one-shot optimal contracts described in Section 3 and all omitted proofs.

2 The model

We consider a two-period–$t \in \{1, 2\}$–environment in which there is a *Limited Partner* (*LP*, the Principal), who has capital to invest in a large number of funds, and a continuum of penniless *General Partners* (*GPs*, the Agents), who have investment opportunities. The LP and each of the GPs may engage in a private equity partnership, also referred to as *the fund*, which is managed by the GP. We consider two dimensions of asymmetric information which, for tractability, we assume to be independent. The first source of asymmetric information is adverse selection. The GP can be either *baseline (b)* or *extraordinary (e)*. As we formalize below, extraordinary GPs deliver higher returns than baseline GPs, although baseline GPs deliver positive profits as well, and thus the LP prefers to engage in a partnership with a baseline GP than leaving its funds idle obtaining a zero return. The second source of asymmetric information is moral hazard. The GP may either *run the fund diligently* and find good investment opportunities, or adopt a passive strategy and generate a lower return.

The fund investment yields a stochastic return whose probability distribution depends on the GP's diligence in managing the fund, which is not verifiable. If the fund is run passively, an investment of k units yields a low return $R_L(k)$ for certain. If the fund is run diligently, investing k units yields a high return $R_H(k)$ with probability $q \in (0, 1)$, and a low return $R_L(k)$ with the remaining probability, where $R_H(k) > R_L(k)$ for all all $k > 0$. We assume that funds require a positive investment to operate, that is, $R_\sigma(0) = 0$, for $\sigma \in \{H, L\}$, and that the return functions are increasing and concave in the fund scale k, that is, $R'_\sigma(\cdot) > 0$, $R''_\sigma(\cdot) < 0$ for all $k \geq 0$. In order to have interior solutions, we also assume that the following Inada conditions hold: $\lim_{k\to 0} R'_\sigma(k) = +\infty$ and $\lim_{k\to 0} R'_\sigma(k)\, k = 0$. Running a fund of size k diligently entails a non-pecuniary cost of $c(k)$ on top of other management costs (which we specify below), with $c(0) = 0$, $c'(\cdot) > 0$ and $c''(\cdot) \geq 0$.[9] The presence of moral hazard affects the structure of payments to incentivize the GP to run the fund diligently.

For notational convenience, we define the expected return from running the fund diligently as $R(k) \equiv qR_H(k) + (1-q)R_L(k)$. We will also refer extensively to the function $\hat{R}(k) \equiv R(k) - c(k)$, which substracts the cost exerted by the GP from the expected return from running the fund diligently. Notice that both $R(\cdot)$ and $\hat{R}(\cdot)$ inherit the monotonicity and curvature properties of $R_H(\cdot)$ and $R_L(\cdot)$. Therefore, both $R(\cdot)$ and $\hat{R}(\cdot)$ are strictly increasing, strictly concave, and satisfy the Inada conditions referred to above. We assume that running the fund diligently is efficient, that is:

Assumption 1 (Efficient to run the fund diligently) $\hat{R}(k) > R_L(k)$.

The GP can be either baseline (b) or extraordinary (e). The type of the GP is its own private information, but it is common knowledge that there is a fraction υ_i of each type, with $\upsilon_b + \upsilon_e = 1$, for $i \in \{b, e\}$. A type$-i$ GP running the fund diligently generates an expected aggregate surplus of:

$$\pi_i(k) \equiv \hat{R}(k) - \theta_i k,$$

with $\theta_b > \theta_e \geq 0$. Hence, extraordinary GPs generate higher profits than baseline GPs, that is, $\pi_e(k) > \pi_b(k) > 0$ for any positive investment scale $k > 0$.[10] We can interpret the term $\theta_i k$, which we assume to be linear for tractability, as the managerial cost that the GP incurs in running the fund. Hence, baseline GPs must incur a higher cost for generating the same investment opportunities as extraordinary GPs. As we shall see below, the adverse selection problem shapes the structure of fees to account for the quality of the GP.

Second-period payoffs are "discounted" by a *duration factor* $\delta \geq 0$ which may be larger

[9] We assume that the cost is non-pecuniary for expositional purposes only. Assuming that the cost is non-pecuniary allows us to abstracting from the interaction between incentives to run the fund diligently and issues of limited liability, which would complicate the analysis without bringing further insight.

[10] As we show in Appendix A.2.4, the assumptions $\lim_{k\to 0} \hat{R}'(k)\, k = 0$ and $\hat{R}(0) = 0$ ensure that $\pi_b(k) > 0$, and thus the baseline GP generate positive profits for any fund size.

than 1 to represent cases where the second period lasts much longer than the first period.[11] In order to shorten notation in some expressions, we write δ_t to represent period$-t$ discount factor, where $\delta_1 = 1$ and $\delta_2 = \delta$. The duration factor determines whether the optimal contract entails types pooling or separation, as we shall see below.

The LP's outside option is normalized to zero in the first period and takes a net value of I in the second period. I represents the opportunity cost of investing in the fund because an alternative investment opportunity with net present value I is missed. The value of I is the realization of a random variable distributed with a common knowledge cumulative distribution function F, positive density f in the whole of its support $[0, +\infty)$, and expected value $\mu \equiv \int_0^{+\infty} IdF(I)$.[12] The realization of I takes place at the beginning of the second period and is only observed by the GP. Since we focus on the impact of the changing investment opportunities on the side of the LP we assume that the GP's outside option is zero in both periods. The LP's second-period uncertain outside option gives raise to the need of finding an efficient means of breaking the fund partnership in case a better investment opportunity arises.

The LP may miss a capital call to enjoy its outside option, but in this case it must pay a pre-specified *default penalty* P_i to the GP. LPs can easily be required to pay a default penalty if they choose to opt out of the fund, as their role in the partnerships consists of providing capital to invest. On the contrary, it seems implausible for an LP to force a GP to run a fund, even if contractually obliged, when the GP could be better off outside the partnership. The threat of mismanaging the fund should suffice to convince the LP to let the GP opt out of the fund. Consistent with this, in reality, penalties–if any–are borne by LPs (Litvak (2004)). Hence, we focus on contracts in which penalties can only be imposed on LPs.

We consider two-period contracts without commitment. Therefore, contracts must be

[11] For instance, in a standard 10-year private equity fund, most of the capital drawdowns occur during the first 5-6 years of the fund life. Early capital calls take place when most of the fund life is ahead.

[12] The unbounded support assumption does not have any implications for the purpose of our analysis and is just made for simplicity of exposition. All we need is that the LP's second-period outside option I takes a sufficiently high value, so that the LP is better off breaking the partnership in some positive probability events.

renegotiation-proof, as any contract subject to a Pareto improvement can be freely terminated and renegotiated in the interim. We assume that partnership agreements are designed by the LP and come as a take-it-or-leave-it offer to GPs.[13] A contract $C_i = \{k_{1i}, k_{2i}, x_{1i\sigma}, x_{2i\sigma}, P_i\}$ designed for a type$-i$ GP specifies, for each round $t \in \{1, 2\}$, the LP's capital contribution $k_{ti} \geq 0$ to the fund, a fee $x_{ti\sigma} \geq 0$ to be paid to the GP when the fund delivers either a high ($\sigma = H$) or a low ($\sigma = L$) return, and a default penalty $P_i \geq 0$ that the LP must pay to the GP if it misses the second-period capital call.[14] This specification potentially allows for pooling contracts, that is, menus of contracts such that $C_b = C_e$. Throughout the text, we call *optimal* contracts to those that maximize the LP's profits. For notational convenience, we write the expected fee obtained by a type$-i$ GP that runs the fund diligently at $t \in \{1, 2\}$ as $x_{ti} \equiv qx_{tiH} + (1-q)\, x_{tiL}$.

Give a contract $C_i = \{k_{1i}, k_{2i}, x_{1i\sigma}, x_{2i\sigma}, P_i\}$, and a particular realization I of the LP's second-period outside option, the expected payoff split between the LP and a type$-i$ GP running the fund diligently is given by:

$$\Pi_L(C_i) = \begin{cases} \sum_{t=1,2} \delta_t \left(R(k_{ti}) - x_{ti}\right) & \text{if the LP meets the capital call} \\ R(k_{1i}) - x_{1i} + \delta\left(I - P_i\right) & \text{otherwise} \end{cases},$$

$$\Pi_i(C_i) = \begin{cases} \sum_{t=1,2} \delta_t \left(x_{ti} - \theta_i k_{ti} - c\left(k_{ti}\right)\right) & \text{if the LP meets the capital call} \\ x_{1i} - \theta_i k_{1i} - c\left(k_{1i}\right) + \delta P_i & \text{otherwise} \end{cases}.$$

For notational convenience, we let $\Pi_{tL}(C_i) \equiv R(k_{ti}) - x_{ti}$ and $\Pi_{ti}(C_i) \equiv x_{ti} - \theta_i k_{ti} - c\left(k_{ti}\right)$ denote the LP's and the type$-i$ GP's tth-period payoff, respectively, when the fund is active in

[13] During the last decade, the contracting position of LPs became increasingly stronger due to a widespread use of gatekeepers, the wider role played by institutional investors, and the introduction of standardized sets of principles, such as those proposed by the Institutional Limited Partners Association (ILPA). See Albert J. Hudec "Negotiating Private Equity Fund Terms. The Shifting Balance of Power," Business Law Today, Volume 19, Number 5 May/June 2010; D. Peninon "The GP-LP Relationship: At the Heart of Private Equity." AltAssets, January 22, 2003; and ILPA Private Equity Principles (January 2011), downloadable from the ILPA website.

[14] The family of contracts that we consider restricts P_i and $x_{2i\sigma}$ to potentially only depend on the GP's type, but *not* on whether the fund realization is high or low. It is straightforward to show that allowing P_i or $x_{2i\sigma}$ to depend on the first-period fund outcome does not change the optimal P_i or $x_{2i\sigma}$.

period $t \in \{1,2\}$. Hence, we have that $\Pi_L(C_i) = \sum_{t=1,2} \delta_t \Pi_{tL}(C_i)$ and $\Pi_i(C_i) = \sum_{t=1,2} \delta_t \Pi_{ti}(C_i)$ for partnerships that extend over two periods. We restrict contracts to satisfy a non-negative payoff constraint for GPs in the first period, that is, $\Pi_{1i}(C_i) \geq 0$, for $i \in \{b, e\}$.[15]

We shall decompose the fees paid by the LP to the GP into *management*, *performance* and *extra* fees. We refer to management fees as the part of the fee that compensates the GP for the managerial cost of running the fund. We call performance fees to those whose payment is linked to a high fund return. As we shall see below, performance fees are included in the optimal contract to incentivize the GP to run the fund diligently. We dub extra fees to the remaining part of the fee structure. As we shall see below, extra fees constitute an additional source of earnings for extraordinary GPs beyond management and performance fees, which are perceived by all GPs.

The timing of contracting is as follows. At the beginning of period $t = 1$, the LP offers a menu of contracts $C = \{C_b, C_e\}$. Upon acceptance of the contract terms, a type$-i$ GP manages a fund of size k_{1i}. At the end of period $t = 1$ the first round ends and the GP collects either a fee x_{1iL} or a fee x_{1iH} if the fund has delivered a low return $R_L(k_{1i})$ or a high return $R_H(k_{1i})$, respectively. At the beginning of period $t = 2$, the GP makes a capital call for the committed amount k_{2i}. They GP may either contribute the committed amount or miss the capital call, in which case it must pay the default penalty P_i to the GP. In the event that the LP pays the default penalty, each party can either collect their respective outside option at $t = 2$, or sign a new contract for the remaining period, specifying a new contribution to the fund k'_{2i}, and fees x'_{2iL} and x'_{2iH}. Upon acceptance of either the original (k_{2i}, x_{2iL}, and x_{2iH}) or the new (k'_{2i}, x'_{2iL}, and x'_{2iH}) terms of the contract, a type$-i$ GP manages the fund of the specified size in exchange for the agreed upon fee. At the end of period $t = 2$, the second round ends and payoffs are realized.

[15] Notice that any feasible contract must meet the GP participation constraint, and thus contracts must satisfy an intertemporal non-negativity constraint $\Pi_i(C_i) \geq 0$. The additional restriction that $\Pi_{1i}(C_i) \geq 0$ is consistent with the GP not having funds of her own and being unable to borrow to run the fund in the first period. Notice that this restriction implies that fees must be non-negative.

3 Benchmark one-period contracts

Before proceeding with the analysis of the problem at hand, we start by laying out benchmark one-shot contracts that we will use in the construction of two-period contracts below. We state the main results here and relegate the formal derivation to the appendix.

3.1 First-best one-period contracts

We first characterize the menu of *first-best contracts* $C^* = \{C_b^*, C_e^*\}$, with $C_i^* = \{k_i^*, x_{i\sigma}^*\}$, in which the LP is perfectly informed of the GP's action and type. By Assumption 1, it follows that it is efficient to instruct the GP to run the fund diligently. Consequently, first-best investment levels k_i^* satisfy $\hat{R}'(k_i^*) = \theta_i$, for $i \in \{b, e\}$. The surplus generated by an efficient one-period contract is therefore given by:

$$\pi_i^* \equiv \hat{R}(k_i^*) - \theta_i k_i^*. \tag{1}$$

Since actions are observable and all individuals are risk-neutral, any combination of fees x_{iL}^* and x_{iH}^* such that $qx_{iH}^* + (1-q)\,x_{iL}^* = \theta_i k_i^* + c\,(k^*)$ is optimal, as they compensate the GP for running the fund diligently in expectation. As shown in the appendix, efficient funds run by extraordinary GPs are larger than those run by baseline GPs (i.e., $k_e^* > k_b^* > 0$), and lead to larger profits (i.e., $\pi_e^* > \pi_b^* > 0$).

3.2 Second-best one-period contracts

We now characterize the menu of *second-best one-period contracts* $C^S = \left\{C_b^S, C_e^S\right\}$, with $C_i^S = \left\{k_i^S, x_{i\sigma}^S\right\}$, corresponding to a one-period environment in which the GP's action and type are not observable. Again, from Assumption 1, it follows that it is optimal to induce the GP to run the fund diligently, which requires a compensation based on performance to provide the GP with the appropriate incentives. Letting $x_i \equiv qx_{iH} + (1-q)\,x_{iL}$, the LP thus solves the following optimization problem:

$$\begin{array}{lll}
\max_{\{k_i \geq 0, x_i \geq 0\}_{i\in\{b,e\}}} & \sum\limits_{i=b,e} \upsilon_i \left[\left(R\left(k_i\right) - x_i\right)\right] & \\
s.t & x_i - c\left(k_i\right) \geq x_{iL} & (MHIC_i) \\
 & x_i - \theta_i k_i - c\left(k_i\right) \geq 0 & (PC_i) \\
 & x_i - \theta_i k_i - c\left(k_i\right) \geq x_j - \theta_i k_j - c\left(k_j\right) & (ASIC_i)
\end{array}, \qquad \text{(S)}$$

where $(MHIC_i)$, (PC_i), and $(ASIC_i)$ stand for type$-i$ GP's moral-hazard incentive constraint, participation constraint, and adverse selection incentive constraint, respectively. The $MHIC$ guarantees that the GP runs the fund diligently. The PC ensures that the GP is willing to run the fund at all. The $ASIC$ ensures that the both GP types prefer their own contracts than impersonating the other type.

Whether this program is solved by a pooling or by a separating contract depends on the likelihood that the GP is extraordinary. In order to reduce the number of cases, but without further loss of insight, we assume from now on that separating contracts dominate pooling contracts in one-shot environments.[16]

Assumption 2 (Proportion of Good GPs) $\nu_e \geq \frac{\pi_b^*}{\pi_e^*}$.

The moral hazard incentive constraint binds for both types of the GP, as otherwise the LP would be giving away resources without improving incentives. Hence, we have that $x_i^S = x_{iL}^S + c\left(k_i^S\right)$, or $x_{iH}^S = x_{iL}^S + \frac{c\left(k_i^S\right)}{q}$. As we show in the appendix, the extraordinary type's adverse selection incentive constraint and the baseline type's participation constraints are also binding. As a result, it follows that $x_b^S = \theta_b k_b^S + c\left(k_b^S\right)$, and that $x_e^S = \theta_e k_e^S + c\left(k_e^S\right) + \Delta\theta k_b^S$,

[16] Notwithstanding, the profit-maximizing two-period contract may entail pooling, as we shall see in Section 6.

where $\Delta\theta \equiv \theta_b - \theta_e > 0$. Hence, Problem (S) simplifies to:

$$\max_{\{k_i \geq 0\}_{i \in \{b,e\}}} \upsilon_b \left(\hat{R}(k_b) - \theta_b k_b \right) + \upsilon_e \left(\hat{R}(k_e) - \theta_e k_e - \Delta\theta k_b \right). \tag{S'}$$

Therefore, the optimal investment levels are characterized by the conditions $\hat{R}'(k_e^S) = \theta_e$ and $\hat{R}'(k_b^S) = \theta_b + \frac{\nu_e}{\nu_b}\Delta\theta$, respectively. Hence, the fund size for extraordinary GPs is efficient, that is, $k_e^S = k_e^*$. However, the fund size for baseline GPs is distorted down to $k_b^S < k_b^*$.

Substituting the optimal fund sizes into the above expressions, we have that fees are given by:

$$x_{i\sigma}^S = \begin{cases} x_{bL}^S = \underbrace{\theta_b k_b^S}_{Management\ fees} \\ x_{eL}^S = \underbrace{\theta_e k_e^*}_{Management\ fees} + \underbrace{\Delta\theta k_b^S}_{Extra\ fees} \\ x_{iH}^S = x_{iL}^S + \underbrace{\frac{c\left(k_i^S\right)}{q}}_{Performance\ fees}, \text{ for } i \in \{b, e\} \end{cases}.$$

GPs are compensated for the managerial cost involved in running the fund through management fees. GPs are also incentivized to running the fund diligently through a performance fee that is paid only in case the fund delivers high returns, and that compensates the GPs for the higher cost incurred by running the fund diligently. In addition, extraordinary GPs receive an extra fee that rewards their higher managerial skills, compensating them for their lower management fee per unit of capital. The LP's expected profit from its portfolio of funds is therefore given by:

$$\Pi_L^S \equiv \upsilon_e \left(\pi_e^* - \Delta\theta k_b^S \right) + \upsilon_b \pi_b^S, \tag{2}$$

where $\pi_b^S \equiv \hat{R}(k_b^S) - \theta_b k_b^S$ is the surplus generated by a baseline GP.

Observe that the profit-maximizing contract addresses the ensuing trade-off between efficiency and rents. The optimal menu of contracts requires a downward distortion $k_b^S < k_b^*$ from the optimal investment for the baseline GP to reduce the extra fees that extraordinary types perceive. Baseline GPs do not have incentives to claim to be an extraordinary GP because the gain $\Delta\theta k_b^S$ in terms of higher fees would be outweighed by the additional managerial cost $\Delta\theta k_e^*$ of running the (larger) fund contemplated for extraordinary GPs.

4 Separating contracts with zero default penalties

In order to assess the role of default penalties in PE contracts, in this section we characterize the LP's profit-maximizing menu of *separating contracts with zero default penalties* $C^Z = \left\{C_b^Z, C_e^Z\right\}$, with $C_i^Z = \left\{k_{1i}^Z, k_{2i}^Z, x_{1i\sigma}^Z, x_{2i\sigma}^Z\right\}$ for $i \in \{b, e\}$ and $\sigma \in \{H, L\}$, where positive penalties are not allowed.[17]

We first determine the contractual terms of the second period. Since in separating contracts types are revealed after the first period, contracts establish the efficient investment levels for both types in the second period. Otherwise, there would be room for a mutually beneficial rearrangement of the terms of the contract. Hence, separating contracts with no penalties satisfy the following condition: $k_{2i}^Z = k_i^*$, for $i \in \{b, e\}$, where these efficient levels are as characterized in Section 3.1.[18] The key aspect to recognize here is that, since all private information about types is revealed in the first period, and the zero-penalties restriction implies that the LP can miss a capital call at no cost, the LP extracts the entire surplus from the GP in the second period. Since contracts are renegotiation-proof, the second-period compensation to the GP must be given by $x_{2iL}^Z = \theta_i k_i^*$ and $x_{2iH}^Z = x_{2iL}^Z + \frac{c\left(k_i^*\right)}{q}$, for $i \in \{b, e\}$.[19] Hence, the LP's second-period payoff is given by $\Pi_{2L}\left(C_i^Z\right) = \pi_i^*$. The fact that the partnership may unilaterally be broken by the LP before the second period at no cost, provided that there are no default penalties, implies that the LP meets the capital call only if the value of its outside

[17] By separating contracts we refer to contracts in which each GP type takes a different contract, hence eliciting its type in the first period. In contrast, pooling contracts specify the same contract terms for both GP types. In this case, their type may only be revealed in the second period.

[18] Any contract in which separation of types is achieved in the first period yields the efficient outcomes in the second period. This phenomenon was first coined by Freixas, Guesnerie, and Tirole (1985) as the *racthet effect* in a different context.

[19] Notice that any front-loaded payment can be renegotiated away in the interim period by the GP. Therefore, it is optimal for the LP to back-load the entire second-period management fee.

option does not exceed the expected rents that it can obtain within the partnership, that is, whenever $I \leq \pi_i^*$.

We now characterize the optimal two-period contract. In order to construct the baseline GP's intertemporal adverse selection incentive constraint, observe that a baseline GP is never willing to run a fund of size k_e^* in the second period in exchange for a transfer $\theta_e k_e^*$, as it incurs a net loss of $\Delta\theta k_e^*$. On the other hand, observe that an extraordinary GP can obtain a net payoff of $\delta\Delta\theta k_b^*$ if it claims to be a baseline GP to perceive a higher management fee, and the partnership extends into the second period, an event that occurs with probability $F(\pi_b^*)$. Hence, the LP solves the following problem:

$$\begin{array}{lll} \max_{\{k_{1i}\geq 0, x_{1i}\geq 0\}, i\in\{b,e\}} & \sum_{i=b,e} \upsilon_i \left(R\left(k_{1i}\right) - x_{1i}\right) + \delta\Pi_{2L}^Z & \\ s.t & x_{1i} - c\left(k_{1i}\right) \geq x_{1iL} & (MHIC_i) \\ & x_{1i} - \theta_i k_{1i} - c\left(k_{1i}\right) \geq 0 & (PC_i) \\ & x_{1b} - \theta_b k_{1b} - c\left(k_{1b}\right) \geq x_{1e} - \theta_b k_{1e} - c\left(k_{1e}\right) & (ASIC_b) \\ & x_{1e} - \theta_e k_{1e} - c\left(k_{1e}\right) \geq x_{1b} - \theta_e k_{1b} - c\left(k_{1b}\right) + \delta\Delta\theta F\left(\pi_b^*\right) k_b^* & (ASIC_e) \end{array}, \tag{Z}$$

where

$$\Pi_{2L}^Z \equiv \sum_{i=b,e} \upsilon_i \left(F\left(\pi_i^*\right)\pi_i^* + \int_{\pi_i^*}^{\infty} IdF\left(I\right)\right)$$

stands for the LP's expected profit in the second period, which does not depend on first-period investment scale or fees.

We turn now to the solution of this problem. Since the LP extracts the entire surplus from the GP in the second period, any incentive compatible contract will have to include an up-front extra fee of $\Delta\theta\left(k_{1b} + \delta F\left(\pi_b^*\right)k_b^*\right)$ to the extraordinary GP, in addition to the management and the performance fees.

Suppose first that the baseline GP's adverse selection incentive constraint does not bind.

From the analysis carried out in Section 3.2, it follows that both the extraordinary GP's adverse selection incentive constraint and the baseline GP's participation constraint bind. Hence, Problem (Z) turns into the following:

$$\max_{\{k_{1i}\}_{i\in\{b,e\}}} \upsilon_b \left(\hat{R}\left(k_{1b}\right) - \theta_b k_{1b}\right) + \upsilon_e \left(\hat{R}\left(k_{1e}\right) - \theta_e k_{1e} - \Delta\theta\left(k_{1b} + \delta F\left(\pi_b^*\right) k_b^*\right)\right) + \delta \Pi_{2L}^Z. \quad \text{(Z')}$$

Therefore, the optimal menu of contracts consists of the same first-period investments as in the one-period second-best contract, that is, k_b^S and k_e^*, as this maximization problem is equivalent to Problem (S'). We label the menu of contracts that solves this program the *no-distortions separating contracts (ND)*, which we will refer to extensively throughout the paper. This menu of contracts is composed of the second-best one-period investment levels in the first period, and the first-best one-period investment levels in the second period.

Definition 1 (No-distortions separating contracts) *The menu of* no-distortions separating contracts $C^{ND} = \left\{C_b^{ND}, C_e^{ND}\right\}$*, with* $C_i^{ND} = \left\{k_{1i}^{ND}, k_{2i}^{ND}, x_{1i\sigma}^{ND}, x_{2i\sigma}^{ND}\right\}$ *for* $i \in \{b,e\}$ *and* $\sigma \in \{H, L\}$*, consists of the second-best one-period investments in the first period (i.e.,* $k_{1b}^{ND} = k_b^S$ *and* $k_{1e}^{ND} = k_e^*$*), efficient investments in the second period (i.e.,* $k_{2b}^{ND} = k_b^*$ *and* $k_{2e}^{ND} = k_e^*$*), and fees* $x_{1bL}^{ND} = \theta_b k_b^S$*,* $x_{1eL}^{ND} = \theta_e k_e^* + \Delta\theta\left(k_b^S + \delta F\left(\pi_b^*\right) k_b^*\right)$*,* $x_{2iL}^{ND} = \theta_i k_i^*$*, and* $x_{tiH}^{ND} = x_{2iL}^{ND} + \frac{c\left(k_{ti}^{ND}\right)}{q}$*, for* $i \in \{b,e\}$ *, and* $t \in \{1,2\}$*.*

However, this contract may not be incentive-compatible because baseline GPs may find it beneficial to pretend they are extraordinary GPs. Observe that a baseline GP obtains a zero payoff by taking its own contract, as its participation constraint binds. But it may pretend to be an extraordinary GP and obtain a net expected payoff of $\Pi_b\left(C_e^{ND}\right) \equiv \Delta\theta\left(k_b^S + \delta F\left(\pi_b^*\right) k_b^* - k_e^*\right)$. If $\Pi_b\left(C_e^{ND}\right) > 0$, the high compensation package that an extraordinary GP receives leads baseline GPs to impersonate extraordinary GPs. In the following lemma we characterize the *incentive-compatibility condition* for any contract with zero penalties.

Lemma 1 (Incentive compatibility) *A pair of investment scales* (k_{1b}, k_{1e}) *involving separation of types in the first period is incentive compatible if and only if:*

$$k_{1e} - k_{1b} \geq \delta F(\pi_b^*) k_b^*. \quad (3)$$

Proof. See Appendix. ■

This condition prescribes that the spread between investment levels for each type must be sufficiently large, so that it is too costly for a baseline type to run too large a fund. The following corollary is a straightforward implication of Lemma 1.

Corollary 1 (Range of no-distortions with zero default penalties) *Define the threshold* $F_b^Z \equiv \frac{k_e^* - k_b^S}{\delta k_b^*}$. *The no-distortions separating contracts menu* C^{ND} *is incentive-compatible if and only if:*

$$F(\pi_b^*) \leq F_b^Z. \quad (4)$$

The threshold F_b^Z establishes a maximum value for the probability $F(\pi_b^*)$ so that the no-distortions contract C^{ND} is incentive-compatible. When Condition (4) holds, the probability of continuation in a fund run by a baseline GP is small enough so that the extra rent that needs to be paid to the extraordinary GP is not attractive enough for the baseline GP to pretend to be a extraordinary one. In this case, the no-distortions separating menu of contracts is optimal.

In what follows, we show that first-period investments have to be distorted away from the one-period second-best investment levels whenever Condition (4) is not satisfied. In this case, the minimal distortion occurs when the baseline GP's adverse selection incentive constraint binds, that is, when the following condition is met:

$$k_{1e}^Z = k_{1b}^Z + \delta F(\pi_b^*) k_b^*. \quad (5)$$

Substituting Expression (5) into the extraordinary GP's adverse selection incentive constraint, we have that $x_{1eL} = \theta_b k_{1e}$. Therefore, we can write Problem (Z) as:

$$\max_{k_{1b} \geq 0} \upsilon_b \left(\hat{R}\left(k_{1b}\right) - \theta_b k_{1b} \right) + \upsilon_e \left(\hat{R}\left(k_{1b} + \delta F\left(\pi_b^*\right) k_b^*\right) - \theta_e (k_{1b} + \delta F\left(\pi_b^*\right) k_b^* \right) + \delta \Pi_{2L}^Z. \qquad \text{(Z”)}$$

Hence, differentiating Expression (Z”) with respect to k_{1b}, it follows that the optimal investment scales satisfy $\upsilon_b \hat{R}'\left(k_{1b}^Z\right) + \upsilon_e \hat{R}'(k_{1e}^Z) = \theta_b$. Notice that the conditions for the optimal investment scales in the one-period separating contract are $\hat{R}'\left(k_e^*\right) = \theta_e$ and $\hat{R}'(k_b^S) = \theta_b + \frac{\nu_e}{\nu_b}\Delta\theta$. Therefore, we have that $\upsilon_b \hat{R}'(k_b^S) + \upsilon_e \hat{R}'\left(k_e^*\right) = \theta_b$. Combining these two observations, it follows that the optimal distortion verifies:

$$\upsilon_b \hat{R}'\left(k_{1b}^Z\right) + \upsilon_e \hat{R}'(k_{1e}^Z) = \upsilon_b \hat{R}'\left(k_b^S\right) + \upsilon_e \hat{R}'\left(k_e^*\right) = \theta_b, \qquad (6)$$

that is, the optimal distortion requires that the expected marginal return of the fund equals that of the one-period profit-maximizing contract. The following proposition summarizes these findings.

Proposition 1 (Optimal separating contracts with zero default penalties) *The optimal menu of separating contracts with zero default penalties is as follows:*

(i) If $F\left(\pi_b^*\right) \leq F_b^Z$, the no-distortions menu of contracts C^{ND} is optimal. First-period investment scales are given by $k_{1b}^Z = k_b^S$ and $k_{1e}^Z = k_e^*$.

(ii) If $F\left(\pi_b^*\right) > F_b^Z$, the no-distortions menu of contracts C^{ND} is not incentive-compatible. The optimal first-period investments k_{1e}^Z and k_{1b}^Z satisfy Conditions (5) and (6), which implies both an upward and a downward distortion of the extraordinary and the baseline GP's investment levels, respectively, that is:

$$k_{1b}^Z < k_b^S < k_e^* < k_{1e}^Z.$$

(iii) The second-period investments are efficient, i.e., $k_{2i}^Z = k_i^*$.

(iv) Fees are given by:

$$x^Z_{ti\sigma} = \begin{cases} x^Z_{1bL} = \underbrace{\theta_b k^Z_{1b}}_{\text{Management fees}} \\ x^Z_{1eL} = \underbrace{\theta_e k^Z_{1e}}_{\text{Management fees}} + \underbrace{\Delta\theta\left(k^Z_{1b} + \delta F\left(\pi^*_b\right) k^*_b\right)}_{\Pi^Z_e:\ \text{Extra fees}} \\ x^Z_{2iL} = \underbrace{\theta_i k^*_i}_{\text{Management fees}} \quad , \textit{for } i \in \{b, e\} \\ x^Z_{tiH} = x^Z_{tiL} + \underbrace{\frac{c\left(k^Z_{ti}\right)}{q}}_{\text{Performance fees}} \quad , \textit{for } i \in \{b, e\},\ t \in \{1, 2\} \end{cases} .$$

(v) The LP meets capital calls efficiently, that is, if and only if $I \leq \pi^*_i$.

The intuition behind this result goes as follows. Since the LP can interrupt the partnership without incurring any cost, the LP would renegotiate any contract to achieve the efficient investment levels and extract the entire surplus created by the partnership in the second period. Anticipating this, an extraordinary GP demands an up-front payment Π^Z_e in the form of extra fees on top of the management and performance fees. By mimicking an extraordinary GP, a baseline GP can obtain a net payoff of $\Pi^Z_e - \Delta\theta k^Z_{1e}$, which is non-positive if and only if $F\left(\pi^*_b\right) \leq F^Z_b$. Hence, whenever $F\left(\pi^*_b\right) > F^Z_b$, the first-period investment levels k^Z_{1b} and k^Z_{1e} must be distorted so that $\Pi^Z_e - \Delta\theta k^Z_{1e} = 0$. Notice that the optimal distortion entails both an upward shift of the extraordinary GP's fund ($k^Z_{1e} > k^*_e$) and a downward shift of the baseline GP's fund ($k^Z_{1b} < k^S_b$), since small distortions around the menu of no-distortions separating contract have a negligible impact on the LP's profit. Partnership interruptions are efficient because the LP extracts the entire second-period surplus π^*_i.

5 Separating contracts with optimally-set default penalties

In this section we characterize the menu of *separating contracts with optimally-set default penalties* $C^D = \left\{C_b^D, C_e^D\right\}$, with $C_i^D = \left\{k_{1i}^D, k_{2i}^D, x_{1i\sigma}^D, x_{2i\sigma}^D, P_i^D\right\}$ for $i \in \{b, e\}$ and $\sigma \in \{H, L\}$, and compare it with the menu of contracts with zero default penalties analyzed in the previous section. We proceed as follows. We first characterize the optimal default penalty for missing a capital call, and the conditions under which the capital call is missed. Then we address the conditions under which contracts with the optimal default penalty need not distort first-period investment scales. As we shall see, the use of an optimally-set default penalty reduces the need for distortions. Finally, we show that distortions, whenever needed, are always strictly smaller than in contracts with a zero default penalty. In Section 6, we characterize the optimal pooling contracts and establish the conditions under which separating contracts dominate pooling contracts.

5.1 Optimal default penalty and partnership termination

We first characterize the optimal default penalty. Let $\Pi_{ti}^D \equiv \Pi_{ti}\left(C_i^D\right)$ denote the tth-period payoff that a type$-i$ GP obtains if it takes its own contract. Analogously, let Π_i^D stand for type$-i$ GP's intertemporal payoff from taking its own contract, and thus:

$$\Pi_i^D \equiv \sum_{t=1,2} \delta_t \Pi_{ti}^D. \tag{7}$$

Again, renegotiation-proof optimal contracts entail efficient investment scales in the second period, that is, $k_{2i}^D = k_i^*$. Hence, if the partnership is not interrupted in the interim, a contract establishing a fee of x_{2i}^D leads to a second-period payoff for a type$-i$ GP of:

$$\Pi_{2i}^D \equiv x_{2i}^D - \theta_i k_i^* - c\left(k_i^*\right). \tag{8}$$

Notice that, unlike the case in which positive default penalties are not allowed, the value of Π_{2i}^D displayed in Expression (8) may be strictly positive. The following proposition characterizes the optimal default penalty, which is determined by the GP's second-period payoff.

Proposition 2 (Optimal default penalty) *The optimal default penalty P_i that the LP must pay to miss the second-period capital call in a partnership with a type$-i$ GP equals the GP's second-period payoff in case the partnership is not interrupted, that is:*

$$P_i = \Pi_{2i}^D. \tag{9}$$

The proof of this result relies on contracts having to be renegotiation-proof. If $P_i < \Pi_{2i}^D$, then the LP could pay the penalty to terminate the contract and renegotiate its terms to $x_{2i}^* = \theta_i k_i^*$, effectively reducing the GP's second-period payoff to P_i. If, on the contrary, $P_i > \Pi_{2i}^D$, the LP could make a take-it-or-leave-it offer $P_i' \in \left(\Pi_{2i}^D, P_i\right)$ to miss the capital call, which the GP would accept, since rejecting the offer would yield a (smaller) rent of Π_{2i}^D.

The following result is a straightforward implication of Proposition 2.

Corollary 2 (Second-period payoff guarantee) *The GP gets the same second-period payoff regardless of whether the LP meets the second-period capital call or not.*

Observe that the LP obtains a second-period payoff of $\pi_i^* - \Pi_{2i}^D$ by meeting the capital call. Interrupting the partnership to enjoy its outside option I requires a payment of the default penalty P_i to the GP. Hence, since $P_i = \Pi_{2i}^D$, it follows that the LP meets capital calls efficiently in contracts with optimally-set default penalties. The following corollary summarizes this fact.

Corollary 3 (Efficient partnership termination with default penalties) *In a separating contract with optimally-set default penalties the LP meets capital calls efficiently, that is, if and only if $I \leq \pi_i^*$.*

5.2 Optimal separating contracts

We now proceed to the characterization of the optimal contract. As above, the contract must incentivize the GP to run the fund efficiently in both periods, that is:

$$x_{ti} - c\left(k_{ti}\right) \geq x_{tiL}. \qquad (MHIC_i)$$

Since GPs are fully insulated against any realization of the outside option, they obtain a payoff of P_i in the second period if it takes a contract designed for herself. Hence, we can write the type$-i$ GP intertemporal participation constraint as:

$$x_{1i} - \theta_i k_{1i} - c\left(k_{1i}\right) + \delta P_i \geq 0. \qquad (PC_i)$$

In order to construct the intertemporal incentive compatibility constraints, notice that impersonating another type in the first period entails taking a gamble in the second period. Consider the case of a type$-i$ GP impersonating a type$-j$, with $i \neq j$. If the LP misses the capital call, then it gets P_j. But if, on the contrary, the LP decides to continue with the partnership, then a type$-i$ GP obtains an expected payoff of $x_{2j}^D - \theta_i k_j^* - c\left(k_j^*\right)$, if it decides to stay in the partnership, or a payoff of 0, in case it opts out of the partnership after the first period. Type$-i$ adverse selection incentive constraint then reads as follows:

$$\begin{aligned} & x_{1i} - \theta_i k_{1i} - c\left(k_{1i}\right) + \delta P_i \geq \\ & x_{1j} - \theta_i k_{1j} - c\left(k_{1j}\right) + \delta\left[F\left(\pi_j^*\right)\max\left\{x_{2j} - \theta_i k_j^* - c\left(k_j^*\right), 0\right\} + \left(1 - F\left(\pi_j^*\right)\right)P_j\right] \end{aligned}. \qquad (ASIC_i)$$

We can then write the LP's problem at time $t=0$ as:

$$\max_{\substack{\{x_{ti}\geq 0, k_{1i}\geq 0\}_{t\in\{1,2\};} \\ i\in\{b,e\}}} \sum_{i=b,e} \upsilon_i \left(\begin{array}{c} \left(\hat{R}\left(k_{1i}\right)-x_{1i}\right)+ \\ +\delta\left[F\left(\pi_i^*\right)\left(\hat{R}\left(k_i^*\right)-x_{2i}\right)+\int_{\pi_i^*}^{\infty}\left(I-P_i\right)dF\left(I\right)\right] \end{array} \right),$$

$$s.t.(MHIC_i),\ (PC_i),\ (ASIC_i),\ \Pi_{1i}\left(C\right)\geq 0 \text{ and } P_i = x_{2i}^D-\theta_i k_i^*-c\left(k_i^*\right) \tag{D}$$

where the last equality follows from Equation (8) and Proposition 2. Solving this problem entails choosing first-period investment levels, as well as first- and second-period fees. Notice that renegotiation-proofness requires second-period investment levels to be efficient. However, unlike the case of contracts with zero default penalties analyzed in the previous section, in which extra fees are front-loaded to the first-period, the optimal contract may (and will) include second-period *extra* fees.

We solve the contracting problem following the same three steps as in the case of zero penalties. We first assume that the baseline GP's adverse selection incentive constraint does not bind and solve for the optimal contract in this case. We then show the conditions under which the baseline GP's adverse selection incentive constraint does not bind. Finally, we characterize the optimal contract when the baseline GP's adverse selection incentive constraint binds.

Suppose first that the baseline GP's adverse selection incentive constraint does not bind. Notice that the baseline GPs' participation constraint must bind, as otherwise the LP would be giving away rents and hardening types separation. Therefore, it follows from Proposition 2 that the default penalty with a baseline type is $P_b=\Pi_b^D=0$. Then, using Equation (8) and Proposition 2, it follows that the baseline GP's second-period management fee is given by $x_{2bL}^D=\theta_b k_b^*$. Using the baseline GP's intertemporal participation constraint displayed in Equation (PC_i), we can also obtain the baseline GP's first-period fees, which are given by $x_{1bL}^D=\theta_b k_{1b}^D$.

Suppose now that all the extraordinary GP's extra fees are paid in full in the second period (we will come back to this point later), i.e., $\Pi_e^D = \delta \Pi_{2e}^D$. In this case, the fees perceived by the extraordinary GP are given by $x_{1e}^D = \theta_e k_{1e}^D + c\left(k_{1e}^D\right)$ and $x_{2e}^D = \theta_e k_e^* + c\left(k_e^*\right) + \Pi_{2e}^D$, where $\Pi_e^D = \Delta\theta\left(k_{1b}^D + \delta F\left(\pi_b^*\right) k_b^*\right)$ is given by the extraordinary GP's adverse selection incentive constraint (Expression $(ASIC_i)$ above), and the fact that partnership interruptions are efficient (Corollary 3). Recognizing that $P_e = \Pi_{2e}^D$, we can rewrite Problem (D) as:

$$\max_{\{k_{1i}\geq 0\}_{i\in\{b,e\}}} \sum_{i=b,e} \upsilon_i \left(\left(\hat{R}\left(k_{1i}\right) - \theta_i k_{1i}\right) - \Pi_i^D + \delta\left(F\left(\pi_i^*\right)\pi_i^* + \int_{\pi_i^*}^{\infty} IdF\left(I\right)\right)\right). \tag{D'}$$

Just as Problem (Z'), this maximization problem is equivalent to Problem (S'). Hence, first-period investment scales are given by $k_{1b}^D = k_b^S$ and $k_{1e}^D = k_e^*$, respectively. Hence, as long as baseline GPs do not want to impersonate extraordinary GP, the no-distortions separating contract C^{ND} defined above is optimal.

We proceed now to addressing the conditions under which the contract menu C^{ND} is incentive compatible. Using the fact that $P_e = \Pi_{2e}^D$, it follows that the baseline GP does not want to impersonate the extraordinary GP if and only if:

$$\Delta\theta k_e^* \geq \delta\left[F\left(\pi_e^*\right)\max\left\{P_e - \Delta\theta k_e^*, 0\right\} + \left(1 - F\left(\pi_e^*\right)\right)P_e\right]. \tag{10}$$

The left hand side of Expression (10) captures the loss from impersonating an extraordinary GP, which stems from the non-compensated portion of the management cost of running a fund intended for an extraordinary GP in the first period. The right-hand side of that expression reflects the gain from impersonating an extraordinary GP, which comes in the second period. On the one hand, with probability $F\left(\pi_e^*\right)$ the LP will honor the capital call. In this case, the baseline GP is willing to continue running the fund if and only if the second-period fees x_{2e} are sufficiently high so that $P_e - \Delta\theta k_e^* \geq 0$. On the other hand, with probability $1 - F\left(\pi_e^*\right)$ the LP will miss the capital call and pay the penalty P_e to the GP.

Expression (10) can therefore take two forms. Suppose first that it is optimal for the baseline GP to run the fund intended for the extraordinary GP in the second period, that is, $P_e - \Delta\theta k_e^* \geq 0$. Rearranging this expression yields $k_{1b} + \delta F(\pi_b^*) k_b^* \geq \delta k_e^*$, in which case Expression (10) reads:

$$k_{1e} + \delta F(\pi_e^*) k_e^* \geq k_{1b} + \delta F(\pi_b^*) k_b^*. \tag{11}$$

Inequality (11) is trivially satisfied by the no-distortion contract menu C^{ND}, since $k_{1b} = k_b^S < k_e^* = k_{1e}$. Therefore, a baseline GP will never want to impersonate an extraordinary GP and run a fund of size k_e^* for two periods.

However, the baseline GP may want to impersonate a extraordinary GP just for the first period, in which case Expression (10) turns into:

$$k_{1e} \geq (1 - F(\pi_e^*))(k_{1b} + \delta F(\pi_b^*) k_b^*). \tag{12}$$

Notice that the only difference between Expressions (3) and (12), which lay out the incentive compatibility conditions for the no-distortion contract menu C^{ND} with zero and with optimally-set default penalties, respectively, is that the right hand side of the latter is multiplied by the probability of partnership termination $1 - F(\pi_e^*) < 1$. Hence, if C^{ND} is incentive compatible with zero default penalties, then C^{ND} is also incentive compatible with optimally-set penalties. Moreover, if the probability of partnership continuation $F(\pi_e^*)$ with an extraordinary GP is sufficiently large, the contract C^{ND} may not incentive compatible with zero penalties and yet be incentive compatible (and thus optimal) with optimally-set penalties. This discussion is formalized in the following proposition. For comparison purposes, Statement (i) in the proposition includes the result laid out in Corollary 1.

Proposition 3 (Distortions on the extensive margin) *Consider the threshold* $F_b^Z \equiv \frac{k_e^* - k_b^S}{\delta k_b^*}$ *defined in Lemma 1, and define the threshold* $F_e^D \equiv 1 - \frac{k_e^*}{k_b^S + \delta F(\pi_b^*) k_b^*}$. *Then:*

(i) The no-distortion menu of contracts C^{ND} *is incentive compatible (and thus optimal) with zero default penalties (i.e.,* $C^Z = C^{ND}$*) if and only if* $F(\pi_b^*) \leq F_b^Z$.

(ii) The no-distortion menu of contracts C^{ND} *is incentive compatible (and thus optimal)*

with optimally-set default penalties (i.e., $C^D = C^{ND}$) if and only if:

$$\text{Either } F(\pi_b^*) \leq F_b^Z \text{ or } \left\{F(\pi_b^*) > F_b^Z \text{ and } F(\pi_e^*) \geq F_e^D\right\}.$$

(iii) If $C^Z = C^{ND}$, then $C^D = C^{ND}$. The converse is not true.

On the *extensive margin*, allowing for default penalties reduces the need of distorting investment scales to the extent that penalties enable the deferral of the extraordinary GP's extra fees to the second period. In contracts with zero default penalties, baseline GPs may contemplate the possibility of impersonating extraordinary GPs to collect large fees in the first period and then walk away from the partnership. In order to prevent this behavior, investment scales must be distorted, as shown in Proposition 1. In contracts with penalties, impersonating an extraordinary GPs entails either bearing the management cost of a large investment scale in both periods, or else only receiving an extra compensation in the form of a default penalty with probability $1 - F(\pi_e^*)$. Hence, baseline GPs have reduced incentives to impersonate extraordinary GPs when there are default penalties to miss capital calls, and thus the reduced need of distortions. As an immediate consequence, when partnerships with extraordinary GPs are sufficiently likely to last for two-periods, that is, for sufficiently large values of $F(\pi_e^*)$, the fund investment scale need not be distorted at all.

A straightforward comparison of Expressions (3) and (12) reveals that the optimally-set penalties also reduce the need for extra distortions on the *intensive margin*. If Expression (12) is not satisfied by the no-distortion contracts menu of contracts C^{ND}, then the baseline GP's adverse selection incentive constraint must bind. Hence, Expression (12) holds with equality, which implies that the spread in first-period investment scales between extraordinary and baseline GPs is smaller than under zero-penalty contracts.

We now proceed to the characterization of the optimal contract. Writing $\Pi_e^D(k_{1b}) \equiv \Delta\theta(k_{1b} + \delta F(\pi_b^*) k_b^*)$ and $\Pi_b^D(k_{1b}) = 0$, with a slight abuse of notation, to recognize the dependence of Π_e^D on the investment scale k_{1b}, and recognizing that $\delta P_e = \Pi_e^D(k_{1b})$, we can write the LP's problem at time $t=0$ when as:

$$\max_{k_{1b}\geq 0} \quad \sum_{i\in\{b,e\}} \upsilon_i \left(\begin{array}{c} \left(\hat{R}\left(k_{1i}\right)-\theta_i k_{1i}\right)-\Pi_i^D\left(k_{1b}\right) \\ +\delta\left(F\left(\pi_i^*\right)\left(\hat{R}\left(k_i^*\right)-\theta_i k_i^*\right)+\int_{\pi_i^*}^{\infty} IdF\left(I\right)\right) \end{array}\right) . \qquad \text{(D”)}$$

$$s.t \qquad k_{1e}=\left(k_{1b}+\delta F\left(\pi_b^*\right)k_b^*\right)\left(1-F\left(\pi_e^*\right)\right)$$

The solution to Problem (D”) is uniquely determined by the following equations. On the one hand, the baseline GP's adverse selection incentive constraint:

$$k_{1e}^D=\left(k_{1b}^D+\delta F\left(\pi_b^*\right)k_b^*\right)\left(1-F\left(\pi_e^*\right)\right). \tag{13}$$

On the other hand, the optimal distortion, which trades off efficiency and rents, takes the form:

$$\upsilon_b\hat{R}'\left(k_{1b}^D\right)+\upsilon_e\hat{R}'\left(k_{1e}^D\right)\left(1-F\left(\pi_e^*\right)\right)=\theta_b-\upsilon_e\theta_e F\left(\pi_e^*\right). \tag{14}$$

Expression (14) is a modified version of Expression (6) that takes into account the reduced likelihood that a baseline GP impersonating an extraordinary GP cashes the extra fees intended for the extraordinary GP. The following proposition characterizes the optimal contract with optimally-set default penalties.

Proposition 4 (Optimal separating contracts with optimally-set default penalties) *The optimal menu of separating contracts with optimally-set default penalties is as follows:*

(i) If either $F\left(\pi_b^*\right)\leq F_b^Z$ *or* $\left\{F\left(\pi_b^*\right)>F_b^Z \text{ and } F\left(\pi_e^*\right)\geq F_e^D\right\}$*, the no-distortions menu of contracts* C^{ND} *is optimal. The first-period investment scales are given by* $k_{1b}^D=k_b^S$ *and* $k_{1e}^D=k_e^*$*.*

(ii) If both $F\left(\pi_b^*\right)>F_b^Z$ *and* $F\left(\pi_e^*\right)<F_e^D$*, the no-distortions menu of contracts* C^{ND} *is not incentive-compatible. The optimal first-period investments* k_{1e}^D *and* k_{1b}^D *satisfy Conditions (13) and (14).*

(iii) Second-period investments are efficient, i.e., $k_{2i}^D=k_i^*$ *for* $i\in\{b,e\}$*.*

(iv) Fees are given by:

$$x_{ti\sigma}^{D}=\begin{cases} x_{1iL}^{D}= & \underbrace{\theta_{i}k_{1i}^{D}}_{\text{Management fees}} \quad \text{, for } i\in\{b,e\} \\ x_{2bL}^{D}= & \underbrace{\theta_{b}k_{b}^{*}}_{\text{Management fees}} \\ x_{2eL}^{D}= & \underbrace{\theta_{e}k_{e}^{*}}_{\text{Management fees}} + \underbrace{\frac{\Delta\theta}{\delta}\left(k_{1b}^{D}+\delta F\left(\pi_{b}^{*}\right)k_{b}^{*}\right)}_{\text{Extra fees}} \\ x_{tiH}^{D}=x_{tiL}^{D}+ & \underbrace{\frac{c\left(k_{ti}^{D}\right)}{q}}_{\text{Performance fees}} \quad \text{, for } i\in\{b,e\},\ t\in\{1,2\} \end{cases}.$$

(v) The LP meets capital calls efficiently, that is, if and only if $I\leq\pi_{i}^{*}$*. The LP pays a default penalty* $P_{e}=\frac{\Delta\theta}{\delta}\left(k_{1b}^{D}+\delta F\left(\pi_{b}^{*}\right)k_{b}^{*}\right)$ *to extraordinary GPs to miss a capital call and pays no default penalty to baseline GPs, that is ,* $P_{b}=0$*.*

5.3 Investment distortions

We now proceed to the comparison of the investment distortions in contracts with zero default penalties and in contracts with optimally-set default penalties. Recall from the previous section that the no-distortions contract is the optimal contract both with zero penalties and with optimally-set penalties if and only if $F\left(\pi_{b}^{*}\right)\leq F_{b}^{Z}$. In this case, the default penalties do not have any bite in reducing the investment distortions. Henceforth we focus on the case $F\left(\pi_{b}^{*}\right)>F_{b}^{Z}$. The following proposition lays out the comparison.

Proposition 5 (Distortions on the intensive margin) *Assume that* $F\left(\pi_{b}^{*}\right)>F_{b}^{Z}$*. Then, it follows that:*

(i) If $F\left(\pi_{e}^{*}\right)\geq F_{e}^{D}$*, first-period investments in contracts with optimally-set default penalties correspond to the second-best one-period investment levels, while investments in contracts with zero default penalties are distorted away from the second-best one-period investment levels, that*

is:

$$k_{1b}^Z < k_{1b}^D = k_b^S < k_e^* = k_{1e}^D < k_{1e}^Z.$$

(ii) If $F(\pi_e^*) < F_e^D$, first-period investments in contracts with optimally-set default penalties are distorted away from second-best one-period investment levels, albeit less than in contracts with zero default penalties, that is:

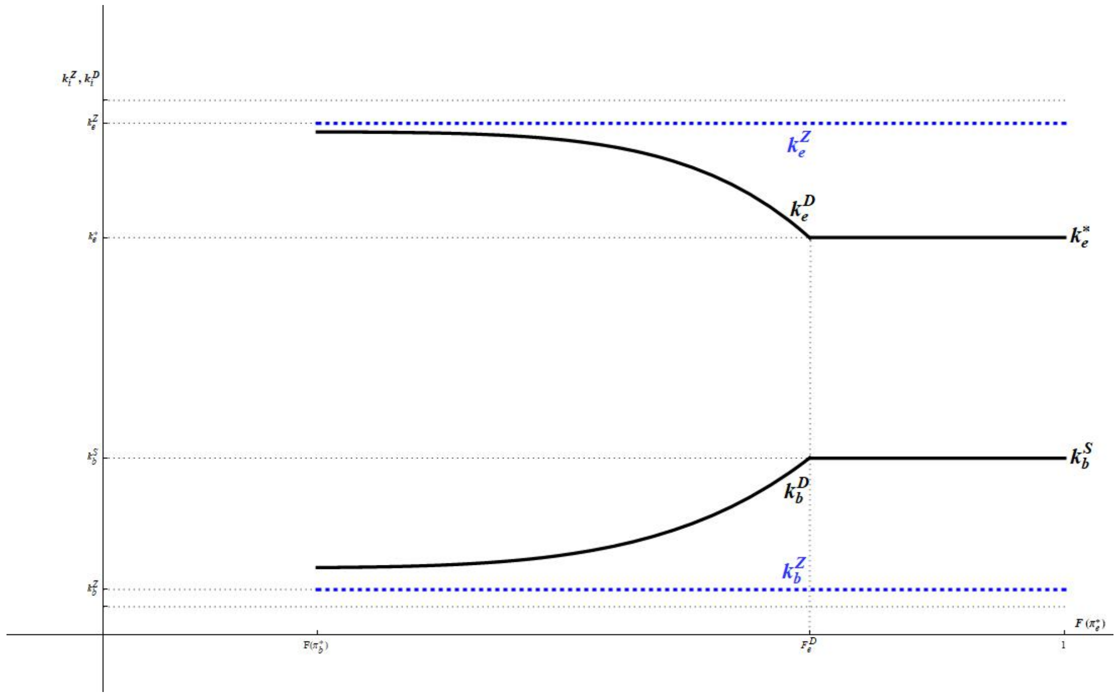
$$k_{1b}^Z < k_{1b}^D < k_b^S < k_e^* < k_{1e}^Z < k_{1e}^D.$$

Moreover, while the distortion scales $k_b^S - k_{1b}^Z$ and $k_{1e}^Z - k_e^*$ are constant in $F(\pi_e^*)$, the distortion scales $k_b^S - k_{1b}^D$ and $k_{1e}^D - k_e^*$ shrink as $F(\pi_e^*)$ increases, getting arbitrarily close to zero as $F(\pi_e^*)$ approaches F_e^D.

Proof. See Appendix. ■

The following figure illustrates the fund size distortions in contracts with zero and with optimally-set default penalties, and how the LP's outside option shapes these distortions. We construct the graph fixing *all-but-one* of the parameters of the model, including $F(\pi_b^*)$, which takes some fix value $F(\pi_b^*) \in (F_b^Z, 1)$. Hence, we depict an environment in which the no-distortion menu of contracts C^{ND} is not incentive compatible in contracts with zero default penalties. We then let $F(\pi_e^*)$ take values in the range $(F(\pi_b^*), 1]$ and plot the first-period fund sizes k_i^Z and k_i^D, for $i \in \{b, e\}$, against different values of $F(\pi_e^*)$.

The dashed blue lines represent the fund sizes k_b^Z and k_e^Z, for the baseline (lower) and the extraordinary (upper) fund sizes, respectively, in contracts with zero default penalties. In this case, the fund sizes are independent of $F(\pi_e^*)$, as stated in Proposition 1. Therefore, the fund sizes are represented by flat lines. The solid black lines represent the fund sizes k_b^D and k_e^D, for the baseline (lower) and the extraordinary (upper) fund sizes, respectively, in contracts with optimally-set default penalties. From Statement (i) in Proposition 5, we know that whenever $F(\pi_e^*) \geq F_e^D$ the no-distortions contract menu is optimal. Therefore, for all the values of $F(\pi_e^*)$ to the right of F_e^D the optimal fund sizes are represented as a flat line corresponding to the values k_b^S and k_e^*, respectively. For values of $F(\pi_e^*)$ to the left of F_e^D the no-distortions contract menu is not incentive compatible and therefore the fund sizes are distorted. As we can observe from Expression (13), and as stated in Proposition 5 (ii), we have that the distortion is larger the lower the probability $F(\pi_e^*)$.



5.4 Optimal deferral of extra fees

Proposition 5 establishes a role for penalties in two-period contracts. In order to preclude a baseline GP from impersonating an extraordinary GP, the LP may defer the payment of the informational rent until the second period, so that the baseline GP has to run a large fund for two periods in order to cash this rent. However, if the LP can miss a capital call without incurring any cost, this deferred payment is not credible: The LP can force a partnership interruption and renegotiate the terms of the agreement. In this case, the LP has to front-load payments in the first period and distort investment scales away from the second-best one-shot contracts in order to make the extraordinary GP's contract unattractive to a baseline GP, as seen in Proposition 1. The role of penalties, by credibly deferring payments to the second period, is that they insulate the extraordinary GP's payoff from partnership terminations without having to make the payment of extra fees up-front, as in the case with zero default penalties. In a certain range, deferring the payment of management fees does, therefore, prevent baseline GPs from mimicking extraordinary GPs without the need of further distortions.

We have assumed above that the payment of extra fees is deferred to the second period. Here, we provide a justification for the reason why postponing part–or all– of the payment of extra fees to the second period is optimal. A glance at Constraint $(ASIC_i)$ above reveals that deferring fees to the second period relaxes the constraint. The left-hand side of the inequality stands for type$-i$ GP's intertemporal payoff, so that any shift of fees across periods, keeping the intertemporal payoff constant, leaves the left hand side unchanged. On the other hand, the right-hand side of the constraint reflects the gains from mimicking the other type. Hence, any impersonation of the other type entailing not running the fund for two periods will be less profitable the higher the amount of extra fees paid in the second period. Hence, deferring extra fees to the second period must be optimal, at least in a weak sense, as it helps incentives. Moreover, notice that deferring payments to GPs does not affect the LP's intertemporal earnings.

In the range in which contracts with zero default penalties need not be distorted there is no need to defer payments to the second period. However, in the range in which including default penalties eliminate distortions, at least a part of the intertemporal informational rent will have to be paid in the second period. For the cases in which even contracts with optimally-set default penalties prescribe some investment distortion, the optimal payment entails deferring the entire extra fee to the second-period. We summarize this discussion in the following proposition.

Proposition 6 (Optimal deferral of extra fees) *Back-loading the payment of any part of the extra fees to the second-period weakly dominates any front-loading payment in the first-period, in the following sense:*

(i)If $F(\pi_b^*) \leq F_b^Z$, the intertemporal allocation of extra fees is irrelevant.

(ii) If both $F(\pi_b^*) > F_b^Z$ and $F(\pi_e^*) \geq F_e^D$, the optimal payment scheme prescribes that at least part of the extra fees are paid in the second period.

(iii) If both $F(\pi_b^*) > F_b^Z$ and $F(\pi_e^*) < F_e^D$, the optimal payment scheme prescribes that the whole extra fees are paid in the second period.

6 Optimal contracts: Separating versus pooling

So far, we have constructed profit-maximizing contracts involving separation of types in the first period. In this section we analyze contracts entailing first-period pooling $C^P = \{C_e^P, C_b^P\}$, with $C^P = \{k_1^P, k_{2i}^P, x_{1\sigma}^P, x_{2i\sigma}^P, P\}$ for $i \in \{b, e\}$ and $\sigma \in \{H, L\}$, and address the effects of the LP's second-period outside option in the design of the optimal contract.

Consider a contract with first-period pooling. Then, the LP's beliefs about the type of the GP are not updated after first-period outcomes are realized. In this case, at the beginning of the second period the contracting parties face a one-period horizon problem. The only renegotiation-proof agreement in which the partnership extends to the second period must therefore be the profit-maximizing one-period contract. Hence, we must have types' separation in the second period, with investment scales of $k_{2b}^P = k_b^S$ and $k_{2e}^P = k_e^*$, respectively. First-period investment levels do not interfere with those of the second period. Hence, the first-period pooling contract must coincide with the one-period optimal pooling contract. Hence, the size of the pooling fund in the first period is $k_1^P = k_b^*$.[20]

Optimal default penalties to miss a capital call must be zero in any contract with types pooling in the first period, that is, $P = 0$. If there is a positive penalty to miss a capital call, it has to be paid to both types of GPs. But that entails a positive transfer to the GP without helping incentives, so that any such penalty strictly reduces the LP's profits. Hence, since the LP can miss a capital call at no cost, the partnership is effectively run by short-term contracts.

We now show that it may be optimal to propose a unique contract in the second period with

[20] Static pooling contracts prescribe that GPs run a fund of size k_b^*. See Appendix A.2.2 for more details.

$k_{2e}^P = k_e^*$, which only extraordinary GPs accept. This contract leads to an expected payoff of $\upsilon_e \pi_e^* + \upsilon_b I$, which exceeds Π_L^S (as defined in Expression (2)) whenever $I > I^P \equiv \pi_b^S - \frac{\upsilon_e}{\upsilon_b} \Delta\theta k_b^S$. The reason why it may be optimal to exclude baseline GPs in the second period in a contract with first-period pooling is the following. When there is separation in the first period, the LP can extract all the second-period surplus. However, if the contract entails pooling in the first period, the LP cannot appropriate the entire second-period surplus, as it must pay an extra fee of $\Delta\theta k_b^S$ to the extraordinary GP to achieve separation, just as in the one-period optimal contract analyzed in Section 3.2. This extra fee reduces the gains from contracting with baseline GPs, which makes the LP's outside option comparatively more valuable. Hence, partnerships with a baseline GP are broken even when it is efficient to continue, namely for realizations $I \in \left(I^P, \pi_b^*\right)$ of the LP's second-period outside option.[21]

The following proposition characterizes the optimal two-period contract with pooling in the first period.

Proposition 7 (Optimal Pooling Contracts) *The optimal pooling contract is as follows:*

(i) Investment is given by $k_1^P = k_b^*$ in the first period.

(ii) Second-period investment scales are second-best efficient, that is, $k_{2b}^P = k_b^S$ and $k_{2e}^P = k_e^*$.

[21] Observe that $I^P > 0$, which follows from the assumption that $\lim_{k \to 0} R'(k)\, k = 0$. Hence, while static profit-maximizing with a zero outside option for the LP prescribes contracting with baseline GPs for certain, it is optimal for the LP to only continue partnerships with extraordinary GPs when his outside option is large enough. Moreover, since $\upsilon_g \pi_g^* + \upsilon_b I^P = \Pi_L^S$, by construction, it follows that $I^P < \pi_b^*$.

(iii) Fees are given by:

$$x^P_{ti\sigma} = \begin{cases} x^P_{1iL} = \underbrace{\theta_b k^*_b}_{\text{Management fees}} \quad \text{, for } i \in \{b,e\} \\ x^P_{2eL} = \begin{cases} \underbrace{\theta_e k^*_e}_{\text{Management fees}} + \underbrace{\Delta\theta k^S_b}_{\text{Extra fees}} & \text{if } I < I^P \\ \underbrace{\theta_e k^*_e}_{\text{Management fees}} & \text{if } I \in \left[I^P, \pi^*_e\right] \end{cases} \\ x^P_{2bL} = \underbrace{\theta_b k^S_b}_{\text{Management fees}} \\ x^P_{tiH} = x^P_{tiL} + \underbrace{\frac{c\left(k^P_{ti}\right)}{q}}_{\text{Performance fees}} \quad \text{, for } i \in \{b,e\},\ t \in \{1,2\} \end{cases}.$$

(iv) Penalties for missing a capital call are zero, that is, $P = 0$.

We now turn to the question of whether the optimal two-period contract is separating or pooling. First, we argue that separating contracts are preferable to pooling contracts as long as the second-period duration is not comparatively too large. Second, we argue that separating contracts are relatively more attractive the better the prospects of the LP to enjoy a valuable outside option in the second period.

The main drawback of separating contracts is that following separation in the first period, the ratchet effect leads to efficient investment levels in the second period. This is quite costly for the LP, for it has to give up large (extra) fees to induce separation in the first period. These rents are larger the longer the second-period duration. Therefore, when the second period extends for a sufficient length, the LP is not willing to incur the cost of inducing separation. Instead, it prefers to give up some efficiency in the first period in exchange for a reduction of rents. Hence, there exists a threshold for the second-period duration such that first-period pooling is preferred to first-period separation for any second-period duration exceeding this threshold.

But this duration threshold depends on the LP's outside option. In order to see why, consider a situation in which the outside option prospects are so good so that it is quite likely that the LP misses a capital call and dissolves the partnership. This context is very similar to a one-period contracting framework. Consequently, in this environment first-period separating contracts dominate pooling contracts regardless of the second-period duration. The opposite situation arises if, for instance, the second-period outside option is (close to) zero with high likelihood, and thus the partnership is very likely to last for two periods. Then, a pooling contract is preferred to separation if the second-period duration is long enough.

In order to formalize the effect of the LP's outside option on the optimal contract, consider a family of probability distribution functions $\{F_\alpha(I) : \alpha \in [0,1]\}$ on the LP's outside option with the following properties:

(i) For any $\alpha' < \alpha''$, $F_{\alpha'}(I) < F_{\alpha''}(I)$ for all $I < \pi_b^*$.

(ii) $F_0(I) = 0$ and $F_1(I) = F(\pi_b^*)$ for all $I < \pi_b^*$.

(iii) For any $\alpha \in [0,1]$, $F_\alpha(\pi_b^*) = F(\pi_b^*) \leq 1$.

Property (i) indexes the family of distributions so that a function with a smaller α first-order stochastically dominates a function with a larger α within the range $[0, \pi_b^*)$. Property (ii) sets the boundaries of the family distribution. In the lower extreme ($\alpha = 0$), we have a distribution with zero mass for all $I < \pi_b^*$, and a mass point at $I = \pi_b^*$. On the upper extreme ($\alpha = 1$), we have a distribution with zero mass for all $I \in (0, \pi_b^*]$, and a mass point at $I = 0$.[22] Finally, Property (iii) establishes that the probability that the outside option I takes a value of at most π_b^* is the same across the entire family. Observe that we leave absolute freedom in the structure of the family for $I > \pi_b^*$, as the trade-off separating-versus-pooling is unaffected by outside value realizations above π_b^*.[23] The following proposition formalizes the previous discussion.

[22] Notice that letting $\alpha = 1$ and $F(\pi_b^*) = 1$ would be equivalent to considering that $I = 0$ for certain.

[23] Observe that if we construct the family leaving the upper part of the distribution unchanged, that is,

Proposition 8 (Optimality of separating versus pooling contracts) *(i) There exists a pooling threshold $\delta^P < \infty$ such that the profit-maximizing contract entails first-period separation of types if $\delta \leq \delta^P$, and first-period types pooling if $\delta > \delta^P$.*

(ii) Let δ^P_α be the pooling threshold associated with the distribution $F_\alpha(I)$ from the family described above. Then, for any $\alpha' < \alpha''$, we have that $\delta^P_{\alpha'} < \delta^P_{\alpha''}$, that is, the lower the probability mass below any value smaller than π^*_b (in a first-order stochastic dominance sense) the smaller the second-period duration range for which separation is preferred to pooling.

Proof. See Appendix. ■

The constructive proof can help the intuition behind this result. On the one hand, if $I < I^P$, first-period separation leads to second-period efficient contracts (with large fees), leading to LP's profits given by $\Pi^*_L = \Pi^S_L - \upsilon_e \Delta\theta \left(k^*_b - k^S_b\right)$. In this range, first-period pooling induces second-period second-best one-period contracts. On the other hand, if $I \in \left[I^P, \pi^*_b\right]$, first-period separation also leads to second-period efficient contracts, leading to LP profits of $\upsilon_e \pi^*_e + \upsilon_b I$. Observe that, by construction of I^P, it follows that $\Pi^*_L < \Pi^S_L < \upsilon_e \pi^*_e + \upsilon_b I$ in the range $I \in \left[I^P, \pi^*_b\right]$. Hence, when there is a positive mass of the outside option value below π^*_b, pooling contracts dominate separating contracts whenever the second-period duration is sufficiently large, i.e., we have $\delta^P < \infty$. Observe also that the advantage of pooling versus separating is larger the better the prospects of the LP's second-period outside option. Hence, for distributions that concentrate higher amounts of mass in larger values of the outside option, pooling is relatively more profitable. Hence, if a distribution first-order stochastically dominates another in the range $[0, \pi^*_b)$, then the former will have a smaller associated pooling threshold.

$F_\alpha(I) = F(I)$ for all $I > \pi^*_b$, then we have that a function with a smaller α first-order stochastically dominates a function with a larger α in the entire range. Consequently, a lower α is associated with a higher outside value expectation μ_α. We can also keep the distribution expectation unchanged across the family, i.e., $\mu_\alpha = \mu$ for all α, which require that for any $\alpha' < \alpha''$, $F_{\alpha'}(I) > F_{\alpha''}(I)$ for some range within the interval (π^*_b, ∞).

7 Empirical predictions

We now lay out some of the empirical predictions that the model yields and relate them to the received evidence.

7.1 Fee structure

The optimal contract features management fees that are proportional to the capital under management, performance fees that reward GPs for delivering high returns, and "extra" fees that are non-proportional to the fund capital. The first set of predictions concerns the structure of fees in private equity funds.

Prediction 1 (Management Fees) *Management fees constitute a percentage of the size of the fund that is smaller the larger the fund under management.*

Prediction 1 stems from the fact that baseline and extraordinary GPs are compensated with a fee $\theta_b k_b$ and $\theta_e k_e$, respectively. Given that $\theta_e < \theta_b$, extraordinary GPs should receive a smaller proportional fee. Additionally, given that in equilibrium extraordinary GPs run larger funds, management fees should represent a smaller percentage of the fund size the larger the fund.

There is empirical evidence that this is the case in practice. Legath (2011) report that over the period $2005 - 2010$ management fees were 2.06% for funds with assets under management below \$500 million, 1.40% for funds between \$500 million and \$1 billion, and 1.23% for funds larger than $\$1 billion$. Gompers and Lerner (1999a) compute the size of a fund as the ratio of the capital invested in the fund to the total amount raised by all other funds, and identify three size groups: partnerships i) with a ratio of $0 - 0.2$ percent; ii) with a ratio of $0.2 - 0.7$; iii) with a ratio greater than 0.7. They find that the present value of management fees for each of these classes is respectively, 19.9%, 18.2%, 15.1% of capital under management. Metrick and Yasuda (2010b) provide practitioners' estimates of annual monitoring fees, which vary between 1% and 5%, with smaller companies at the high end of this range and larger companies at the low end. This evidence combined provides support for the model's prediction that large funds receive lower management fees per unit of capital invested.

In our framework, performance fees reward GPs if and only if the fund delivers high profits, and are given by $x_{tiH} - x_{tiL} = \frac{c(k_{ti})}{q}$, thus representing a fraction $\frac{c(k_{ti})}{q(R_H(k_{ti}) - \theta_i k_{ti})}$ of the net fund profit. This fraction is higher the higher the cost of running the fund diligently, which is captured by the cost function $c(\cdot)$. We should expect that the more complex the task of running the fund with diligence, the higher the performance-based reward. Additionally, performance-based fees should increase the lower the likelihood that the funds delivers high profits, as captured by the uncertainty parameter q. These observations motivate the following predictions.

Prediction 2 (Performance Fees) *Performance fees constitute a larger fraction of the fund under management the more uncertain the return of the fund and the more difficult the tasks to manage the fund.*

Toll and Vayner (2012) document that at least one-third of the funds in their sample delivered a carried interest that was considerably below the once-standard 20% benchmark. They report that "Their typically lower-than-20-percent carried interests (or in some cases no carried interest at all) reflect the lower degree of difficulty in their investment strategy, compared with most other private equity firms". Additionally, they argue that "downward pressure is also being applied to carried interest at firms whose investment strategy involves less risk than your run-of-the-mill private equity firm–those managing infrastructure funds, credit opportunity funds and related vehicles".

As discussed in Phalippou, Rauch, and Umber (2018), transaction fees are seldom fully rebated against management fees.[24] Therefore, transaction fees represent charges that GPs

[24] For example, Phalippou *et al.* report that between 2007 and 2012 Apollo rebated 61% of the transaction fees, while KKR rebated 39%. Legath (2011) reports that: "Approximately 43.7% of the private equity firms split the fees evenly between the general partner and/or an affiliated advisory entity and the limited partners. The remaining 19.7% of the firms provide that all or a significant portion of the fees are paid to the general partner of the private equity firm."

impose on the fund and that indirectly reduce the LPs' profits. In our model, LPs pay extra fees to better GPs, who also run larger funds. Extra fees in the model are non-proportional to the fund size. The following prediction is derived from these observations.

Prediction 3 (Extra Fees) *The compensation through extra fees is non-proportional to the capital under management, and constitute a relatively higher source of income for larger funds.*

Metrick and Yasuda (2010b) report that the GPs of buyout funds charge transaction fees that vary between 1% and 2% of the transaction value. Phalippou, Rauch, and Umber (2018) show that transaction fees are an important source of revenue for GPs. Transaction fees are charged in 75% of LBO related deals and represent 0.81% of total enterprise value of the target. The log of the transaction fees increases in the log of total enterprise value. This evidence is supportive of the idea that non-proportional fees are levied by GPs (directly on portfolio firms, and indirectly on LPs), and that transaction fees constitute a higher source of income for larger funds.

7.2 Default Penalties

The second set of predictions have to do with default penalties. In our framework, only extraordinary GPs must be compensated in the form of a default penalty for the loss of (extra) fees when the LP dishonors a capital call (Propositions 4 (v) and 7). Moreover, our model predicts that better GPs run larger funds. The combinations of these facts yields the following prediction.

Prediction 4 (GP quality, fund size, and penalty severity) *The GP quality, the size of the fund, and the severity of the default penalty are positively correlated.*

The only analytical academic work on default penalties is Litvak (2004), which shows that default penalties are higher (in terms of coefficient of severity) in larger funds. Moreover, Litvak (2009) shows that better GPs run larger funds. Insofar as Litvak's coefficient of severity is positively correlated with a monetary loss for the LP, the combined evidence of Litvak (2004) and Litvak (2009) is consistent with our prediction: quality of GPs, size of the fund, and severity of the penalty are positively correlated.

From Proposition 2, it follows that the present value of the optimal default penalty is given by $\Pi_e^D = k_{1b}^D + \delta F(\pi_b^*) k_b^*$, which is increasing in the second-period duration. The following prediction is an immediate consequence of this property.

Prediction 5 (Option term) *Default penalties are larger the longer the partnership duration after the capital call.*

Litvak (2004) shows that default penalties are positively correlated with the option term, which constitutes a measure of the relative importance of future capital calls, thus providing support for this prediction.[25]

In our model, extra fees are intended to reward extraordinary GPs. Default penalties are set in contracts with extraordinary GPs to safeguard their promised compensation. Hence, we should expect that funds that plan to reward GPs through fees other than management and performance fees include default penalties. Moreover, we should expect that the size of the default penalty be similar to the size of the scheduled fees other than management and performance fees to be paid to the GP. The following prediction stems from these observations.

Prediction 6 (Forfeited extra fees) *The size of the default penalty is similar to the size of the forfeited fees, other than management and performance fees, upon missing a capital call*

[25] For instance, suppose that the life of the fund is two years and that a maximum fraction α of the committed capital can be called at the fund's inception. The option term is given by $100/\alpha$. Hence, the option term decreases with the amount of capital that can be called at the fund's inception, being 100 if the entire committed capital is callable at the beginning, and being 200 if only half the commited capital is callable when the fund is started. Therefore, the option term increases with the amount of capital that has to be called in the future.

(in partnerships with zero default penalties the GP only gets compensated through management and performance fees).

Toll and Vayner (2012) provide evidence that around 70% of North American funds include default penalty provisions consisting of the forfeiture of a portion of the capital balance to compensate the GP for its loss. In addition, Litvak (2004) finds that the severity of default penalty provisions is negatively correlated with management fees, which provides indirect evidence that default penalties are larger in funds whose GPs are rewarded through alternative fees. However, to the best of our knowledge, there does not exist any systematic cross-sectional evidence to date relating the size of default penalties and the size of fees that GPs expect to receive in the course of the partnership.

In our framework, default penalties reduce the need for investment distortions in combination with the deferral of extra fees. While management and performance fees are paid on a period-by-period basis, it is optimal that at least a fraction of the extra fees is deferred to the second period (Proposition 6). Hence, we should expect that funds with default penalty clauses defer their fees other than management fees and carried interest, which motivates the following prediction.

Prediction 7 (Deferral of extra fees) *In private equity funds in which default penalties are included, the duration of payment of extra fees (transaction, monitoring, advising,...) is longer than the duration of management fees and carried interest.*

There is ample evidence showing that while management fees are perceived periodically (typically, yearly) from the very inception of funds, extra fees are mostly charged at later times, once investments materialize. However, there is no evidence to date of any kind relating the timing of fees payment with default penalties.

8 Conclusions

This paper provides a framework to study the interaction between LPs and GPs in a two-period model with asymmetric information. LPs set the contractual terms and conditions to screen GPs of heterogeneous and unobservable ability and to provide incentives to run funds diligently. Optimal contracts include default penalties that LPs have to pay if they miss a capital call. The model is particularly suited to analyzing the screening process that LPs go through when selecting GPs with little or no previous history. In such a setting the bargaining power of LPs is particularly strong and adverse selection among GPs is potentially severe.

We show that LPs may distort investment scales to better screen GPs and that the extent of the distortion is smaller if default penalties are included in the agreement. Default penalties act as an insurance mechanism for better GPs, and give the contracting parties the ability to defer some of the fees to the second period.

The model draws predictions on the use and the size of default penalties, which are a set of clauses that has received little attention in the academic literature, despite their common usage. As it happens in reality, default penalties are higher in larger funds, and larger funds are run by better-performing GPs. The model also draws predictions on the fee structure. The common fee structure employed by private equity funds comprises three types of fees: management fees, performance fees, and an additional set of (extra) fees, including transaction, advisory or monitoring fees. Our model shows that management fees should represent a smaller percentage in larger funds. The model also draws predictions on fees that are not proportional to the capital under management, but that should be larger for better GPs, and that should be paid not at the inception of the fund, but later in time. The predictions of the model offer possible avenues for future empirical research in the field of private equity contracts.

A Derivation of benchmark contracts

In this section, we provide a derivation of the benchmark contracts that have been laid out in Section 3.

A.1 Contracts with perfect information

Suppose first that the GP's type is known to the LP. The efficient investment level k_i^* for a type$-i$ GP is obtained by equating the marginal return of the investment to the marginal cost incurred by the type$-i$'s marginal cost, that is:

$$\hat{R}'(k_i^*) = \theta_i.$$

Notice that the Inada condition $\hat{R}'(0) = +\infty$ ensures that $k_i^* > 0$. Moreover, we have that $k_b^* < k_e^*$, which follows directly from the concavity of $\hat{R}(k)$ and the fact that $\theta_e < \theta_b$. Hence, it is efficient that an extraordinary GP runs a larger fund.

A type$-i$ GP would then get a transfer of:

$$x_i^* = \theta_i k_i^* + c(k_i^*).$$

The LP would obtain the returns of the fund and, after paying the fees to the GP, it would obtain a payoff of:

$$\pi_i^* \equiv \hat{R}(k_i^*) - \theta_i k_i^*.$$

Notice that the amount π_i^* not only stands for the GP's payoff in a full information regime, but that it also corresponds to the efficient surplus that could be generated by a type$-i$ GP. Notice that $\pi_e^* > \pi_b^*$, that is, an extraordinary GP can potentially generate a higher surplus.[26]

A.2 Second-best one-period contracts

In a one-period framework in which the GP's action and type are not observable, the LP solves the following optimization program:

$$\begin{array}{lll}
\max_{\{k_i \geq 0, x_i \geq 0\}_{i\in\{b,e\}}} & \sum_{i\in\{b,e\}} v_i \left[\left(\hat{R}(k_i) - x_i \right) \right] & \\
s.t & x_i - c(k_i) \geq x_{iL} & (MHIC_i) \\
 & x_i - \theta_i k_i - c(k_i) \geq 0 & (PC_i) \\
 & x_i - \theta_i k_i - c(k_i) \geq x_j - \theta_i k_j - c(k_j) & (ASIC_i)
\end{array},$$

where $(MHIC_i)$, (PC_i), and $(ASIC_i)$ stand for type$-i$ GP's moral-hazard incentive constraint, participation constraint and adverse selection incentive constraint, respectively. Whether this program is solved by a pooling or by a separating contract depends on the ex-ante likelihood that the GP is extraordinary. In what follows, we provide the conditions for one-shot separating equilibria to be preferable than pooling contracts.

[26] By strict concavity of $\hat{R}(k)$, and optimality of the extraordinary GP's efficient investment level, we have that $\pi_e^* \equiv R(k_e^*) - \theta_e \cdot k_e^* > R(k_b^*) - \theta_e \cdot k_b^*$. Also, since $k_b^* > 0$, we have that $\theta_e < \theta_b$ implies that $R(k_b^*) - \theta_e \cdot k_b^* > R(k_b^*) - \theta_b \cdot k_b^* \equiv \pi_b^*$.

A.2.1 One-period separating contracts

First, observe that the baseline GP's participation constraint must bind, as giving up rents to this type would only harden the adverse selection incentive constraint for the extraordinary type. Hence, in a *separating second-best one-period (S)* contract, we have that $x_b^S = \theta_b k_b^S + c\left(k_b^S\right)$. Moreover, the extraordinary GP's adverse selection incentive constraint must bind as well, as otherwise the extraordinary GP would be granted an unnecessarily large rent. Hence, we have that $x_{1e} = \theta_e k_e^S + c\left(k_e^S\right) + \Delta\theta k_b^S$. The LP's problem then reduces to:

$$\max_{\{k_i \geq 0\}_{i \in \{b,e\}}} \quad \sum_{i=b,e} \upsilon_i \left[\left(\hat{R}\left(k_i\right) - \theta_i x_i\right)\right] - \upsilon_e \Delta\theta k_b \quad ,$$

with fees being as specified above.

Since this program is concave, an interior solution to this problem is characterized by its first-order conditions. Then, the separating menu of contracts entails an efficient investment level for the extraordinary type, that is, $k_e^S = k_e^*$ which, as seen above, satisfies:

$$\hat{R}'(k_i^*) = \theta_i.$$

However, the baseline type's investment level satisfies:

$$\hat{R}'(k_b^S) = \theta_b + \frac{\nu_e}{\nu_b}\Delta\theta.$$

Clearly, we have that $k_b^S < k_b^*$, which follows directly from the concavity of $\hat{R}$ and the fact that $\theta_b + \frac{\nu_e}{\nu_b}\Delta\theta > \theta_b$.

A.2.2 One-period pooling contracts

Consider now the option of granting a unique *static pooling contract (SP)*, and thus $k_b = k_e = k^{SP}$ and $x_b = x_e = x^{SP}$. Then, the incentive compatibility constraints are trivially satisfied. Also, since $\theta_e < \theta_b$, the participation constraint for the baseline GP implies that the extraordinary GP's participation constraint is also satisfied. Hence, since giving up rents to the baseline type would be suboptimal, it follows that $x^{SP} = \theta_b k^{SP}$. Therefore, the LP's problem reduces to:

$$\max_{k \geq 0} \quad \hat{R}\left(k\right) - \theta_b k.$$

Since $\hat{R}$ is concave and differentiable, the first-order condition is both necessary and sufficient to characterize an interior solution to this program. The *static pooling (SP)* investment level k^{SP} then satisfies:

$$\hat{R}'(k^{SP}) = \theta_b.$$

Hence, a pooling contract investment level corresponds to the efficient investment level for a baseline GP, that is, $k^{SP} = k_b^* > 0$.

A.2.3 One-period optimal contracts

We now analyze the conditions under which the optimal one-period contract is separating. Offering a pair of contracts with investment levels $\left\{k_b^S, k_e^*\right\}$ grants the LP an expected payoff of:

$$\Pi_L^S \equiv \upsilon_e \left(\pi_e^* - \Delta\theta k_b^S\right) + \upsilon_b \pi_b^S,$$

where $\pi_b^S \equiv R\left(k_b^S\right) - \theta_b k_b^S$ stands for the surplus generated by a baseline GP when it invests the amount k_b^S prescribed by a separating one-period contract.

The LP's payoff from a one-period pooling contract is given by:

$$\Pi_L^{SP} \equiv \hat{R}\left(k_b^*\right) - \theta_b k_b^* \equiv \pi_b^*.$$

Effectively, in a pooling contract the LP gets the same profits as if the GP was known to be the baseline type, for all the extra surplus generated by an extraordinary GP is fully appropriated by the extraordinary GP itself.

Whether the profit-maximizing one-period contract entails pooling or types separation depends on the ex-ante likelihood υ_e that the GP is extraordinary. A menu of separating contracts specifies an efficient investment level k_e^* and a (relatively small) transfer $\Delta\theta k_b^S$ to an extraordinary GP. Pooling contracts, on the contrary, prescribe an inefficiently low level of investment k_b^* for a (relatively large) transfer $\Delta\theta k_b^*$ to an extraordinary GP. Hence, the more likely the GP is extraordinary, the better a separating contract.[27] Assumption 2 guarantees that the optimal one-period contract is separating.

A.2.4 Excluding the baseline type

In this section, we provide an analysis of the baseline type non-exclusion condition. If the LP offers a unique contract specifying $k = k_e^*$, then it obtains a payoff of $\nu_e \pi_e^*$. On the other hand, offering a pair of contracts with investment levels $\left\{k_b^S, k_e^*\right\}$ grants the LP an expected payoff of Π_L^S. A separating menu of contracts dominates a unique separating contract as long as the expected gain $\upsilon_b \pi_b^S$ from potentially contracting with a baseline GP exceeds the expected transfer $\upsilon_e \Delta\theta k_b^S$ needed to induce an extraordinary GP to choose its own contract. We can write this difference as:

$$\begin{aligned}
\upsilon_b \pi_b^S - \upsilon_e \Delta\theta k_b^S &= \upsilon_b \left[\hat{R}\left(k_b^S\right) - \left(\theta_b k_b^S + \frac{\upsilon_e}{\upsilon_b}\Delta\theta\right) k_b^S\right] \\
&= \upsilon_b \left[\hat{R}\left(k_b^S\right) - \hat{R}'(k_b^S) k_b^S\right]
\end{aligned}$$

Observe that $\hat{R}\left(k\right) - \hat{R}'(k)k$ is strictly increasing in k which, coupled with the assumptions that $\lim'_{k\to 0} \hat{R}\left(k\right) k = 0$ and $\hat{R}\left(0\right) = 0$, ensures that $\upsilon_b \pi_b^S - \upsilon_e \Delta\theta k_b^S$ is positive. Nonetheless, although excluding a baseline GP is never optimal in a one-period setting, the optimal two-period contract may entail exclusive contracting with an extraordinary GP, as we have seen above.

B Omitted Proofs

Proof of Lemma 1.

Combining the adverse selection incentive constraints of both types of GPs, we have that a menu of separating contracts is incentive-compatible only if the following conditions are satisfied:

$$\theta_e \left(k_{1e} - k_{1b}\right) + \delta \Delta\theta F\left(\pi_b^*\right) k_b^* \leq x_{1e} - x_{1b} \leq \theta_b \left(k_{1e} - k_{1b}\right).$$

[27] However, a pooling contract may never be profit-maximizing. Although a pooling contract prescribes an efficient level of investment for a baseline GP, it also specifies a large transfer to the extraordinary one. As the likelihood that the GP is a baseline type increases, separating contracts approach pooling contracts, since the investment distortion for the baseline GP gets arbitrarily close to zero. Depending on parameters, we may have that separating contracts approach pooling contracts through a path in which pooling contracts are always dominated.

The incentive-compatibility condition immediately follows from combining the first and the second inequality.

Conversely, if condition (3) holds, then the baseline GP's adverse selection incentive constraint is satisfied, and thus the pair of investment levels k_{1b} and k_{1e} are incentive compatible. ■

Proof of Proposition 5.

(i) Shown in the main text.

(ii) The first derivative (w.r.t. k_{1b}) of the LP's optimization problem is given by:

$$\frac{\partial \Pi_L}{\partial k_{1b}}\left(k_{1b}^D, k_{1e}^D\right) = \upsilon_b \cdot \left(\hat{R}'\left(k_{1b}^D\right) - \theta_b\right) + \upsilon_e \cdot \left(\left(\hat{R}'\left(k_{1e}^D\right) - \theta_e\right) \cdot \left(1 - F\left(\pi_e^*\right)\right) - \Delta\theta\right).$$

Letting $\frac{\partial \Pi_L}{\partial k_{1b}}\left(k_{1b}^D, k_{1e}^D\right) = 0$ yields:

$$\upsilon_b \cdot \hat{R}'\left(k_{1b}^D\right) + \upsilon_e \cdot \hat{R}'\left(k_{1e}^D\right) \cdot \left(1 - F\left(\pi_e^*\right)\right) = \theta_b - \upsilon_e \cdot \theta_e \cdot F\left(\pi_e^*\right).$$

Observe that for $F\left(\pi_e^*\right)$ sufficiently small, the condition reads:

$$\upsilon_b \cdot \hat{R}'\left(k_{1b}^D\right) + \upsilon_e \cdot \hat{R}'\left(k_{1e}^D\right) = \theta_b,$$

which is the condition that determines the distortion for contracts with zero penalties. We have shown in the main text that $k_{1b}^Z < k_b^S < k_e^* < k_{1g}^Z$. We now show that k_{1b}^D approaches k_b^S as $F\left(\pi_e^*\right)$ increases. First, we have that:

$$\frac{\partial^2 \Pi_L}{\partial \left(k_{1b}\right)^2}\left(k_{1b}^D, k_{1e}^D\right) = \upsilon_b \cdot \hat{R}''\left(k_{1b}^D\right) + \upsilon_e \cdot \hat{R}''\left(k_{1e}^D\right) \cdot \left(1 - F\left(\pi_e^*\right)\right) < 0.$$

Also,

$$\frac{\partial^2 \Pi_L}{\partial k_{1b} \partial F\left(\pi_e^*\right)}\left(k_{1b}^D, k_{1e}^D\right) = \left(\begin{array}{c} -\upsilon_e \cdot \left[\left(1 - F\left(\pi_e^*\right)\right) \cdot \hat{R}''\left(k_{1e}^D\right) \cdot f\left(\pi_e^*\right) \cdot \left(k_{1b} + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right)\right] \\ + \left(\hat{R}'\left(k_{1e}^D\right) - \theta_e\right) \cdot f\left(\pi_e^*\right) \end{array}\right) > 0,$$

where the last inequality follows from the fact that $\hat{R}'\left(k_e^*\right) = \theta_e$, $k_{1e}^D > k_e^*$ and $\hat{R}\left(\cdot\right)$ being strictly concave. Hence, it follows that $\frac{dk_{1b}}{dF\left(\pi_e^*\right)} > 0$. By the same token, we have that

$$\frac{\partial \Pi_L}{\partial k_{1e}}\left(k_{1b}^D, k_{1e}^D\right) = \upsilon_b \cdot \left(\hat{R}'\left(k_{1b}^D\right) - \theta_b\right) \cdot \frac{1}{1 - F\left(\pi_e^*\right)} + \upsilon_e \cdot \left(\hat{R}'\left(k_{1e}^D\right) - \theta_e - \Delta\theta \cdot \frac{1}{1 - F\left(\pi_e^*\right)}\right),$$

which leads to:

$$\frac{\partial^2 \Pi_L}{\partial \left(k_{1e}\right)^2}\left(k_{1b}^D, k_{1e}^D\right) = \upsilon_b \cdot \hat{R}''\left(k_{1b}^D\right) \cdot \frac{1}{\left(1 - F\left(\pi_e^*\right)\right)^2} + \upsilon_e \cdot \hat{R}''\left(k_{1e}^D\right) < 0,$$

and

$$\frac{\partial^2 \Pi_L}{\partial k_{1e} \partial F\left(\pi_e^*\right)}\left(k_{1b}^D, k_{1e}^D\right) = \left(\left(\begin{array}{c} \upsilon_b \cdot \hat{R}''\left(k_{1b}^D\right) \cdot \left(k_{1b}^D + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right) + \\ \left[\upsilon_b \cdot \left(\hat{R}'\left(k_{1b}^D\right) - \theta_b\right) - \upsilon_e \cdot \Delta\theta\right] \end{array}\right) \cdot \frac{f\left(\pi_e^*\right)}{\left(1 - F\left(\pi_e^*\right)\right)^2}\right) < 0,$$

where the last inequality follows from the fact that $\upsilon_b \cdot \left(\hat{R}'\left(k_{1b}^D\right) - \theta_b\right) - \upsilon_e \cdot \Delta\theta < 0$ (observe that

$\upsilon_b \cdot \left(\hat{R}'\left(k_b^S\right) - \theta_b\right) - \upsilon_e \cdot \Delta\theta = 0$ and that $\hat{R}'\left(k_b^S\right)$ is strictly increasing) and the strict concavity of $\hat{R}(\cdot)$. Hence, $\frac{dk_{1e}}{dF(\pi_e^*)} < 0$. ■

Proof of Proposition 8.

(i) With first-period pooling, the LP's profits are given by:

$$\pi_b^* + \delta \cdot \upsilon_e \cdot \left(F\left(\pi_e^*\right) \cdot \pi_e^* + \int_{\pi_e^*}^{\infty} I \cdot dF(I)\right) + \delta \cdot \upsilon_b \cdot \left(F\left(I^P\right) \cdot \pi_b^S + \int_{I^P}^{\infty} I \cdot dF(I)\right)$$
$$-\delta \cdot \upsilon_e \cdot \Delta\theta \cdot F\left(I^P\right) \cdot k_b^S.$$

With first-period separation, the LP's profits are bounded above by[28]:

$$\upsilon_e \cdot \pi_e^* + \upsilon_b \cdot \pi_b^S - \upsilon_e \cdot \Delta\theta \cdot \left(k_b^S + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right)$$
$$+\delta \cdot \upsilon_e \cdot \left(F\left(\pi_e^*\right) \cdot \pi_e^* + \int_{\pi_e^*}^{\infty} I \cdot dF(I)\right) + \delta \cdot \upsilon_b \cdot \left(F\left(\pi_b^*\right) \cdot \pi_b^* + \int_{\pi_b^*}^{\infty} I \cdot dF(I)\right).$$

Hence, the difference between first-period separation and first-period pooling is bounded above by:

$$\Pi_L^S - \pi_b^* + \delta \cdot \upsilon_b \cdot \left(\left(F\left(\pi_b^*\right) \cdot \pi_b^* - F\left(I^P\right) \cdot \pi_b^S\right) - \int_{I^P}^{\pi_b^*} I \cdot dF(I)\right) - \delta \cdot \upsilon_e \cdot \Delta\theta \cdot \left(F\left(\pi_b^*\right) \cdot k_b^* - F\left(I^P\right) \cdot k_b^S\right),$$

which we can write as:

$$\left[\Pi_L^S - \pi_b^*\right] + \delta \cdot \left[F\left(I^P\right) \cdot \left(\Pi_L^* - \Pi_L^S\right) + \left(F\left(\pi_b^*\right) - F\left(I^P\right)\right) \cdot \left(\Pi_L^* - \upsilon_e \cdot \pi_e^*\right) - \upsilon_b \cdot \int_{I^P}^{\pi_b^*} I \cdot dF(I)\right], \tag{15}$$

where

$$\Pi_L^* \equiv \upsilon_e \cdot \left(\pi_e^* - \Delta\theta \cdot k_b^*\right) + \upsilon_b \cdot \pi_b^S$$

stands for the LP's profits if the efficient one-period investment scales are implemented and appropriate separation rents are paid to extraordinary GPs. Since the second-best contract yields Π_L^S to the LP and is optimal, it follows that $\Pi_L^* - \Pi_L^S < 0$. Moreover, we know by construction of I^P that $\Pi_L^* < \upsilon_e \cdot \pi_e^* + \upsilon_b \cdot I$ for any $I \in \left(I^P, \pi_b^*\right)$, so that $\Pi_L^* < \upsilon_e \cdot \pi_e^* + \frac{1}{F\left(\pi_b^*\right) - F\left(I^P\right)} \upsilon_b \cdot \int_{I^P}^{\pi_b^*} I \cdot dF(I)$. Hence, while the first bracketed term in expression (15) is positive, the second bracketed term is negative. The first proposition result follows immediately.

(ii) First, observe that by optimality of second-best contracts and by construction of I^P, we have that $\Pi_L^* < \Pi_L^S < \upsilon_e \cdot \pi_e^* + \upsilon_b \cdot I$. Consider $\alpha' < \alpha''$, so that $F_{\alpha'}$ first-order stochastically dominates $F_{\alpha''}$ in the range $[0, \pi_b^*)$. Then, $\Pi_L^* - \Pi_L^S$ carries a relatively lower weight than $\Pi_L^* - \upsilon_e \cdot \pi_e^* + \upsilon_b \cdot I$ in expression (15) under $F_{\alpha'}$ than under $F_{\alpha''}$. Hence, the second bracketed term in expression (15) is smaller (larger in absolute value) under $F_{\alpha'}$ than under $F_{\alpha''}$, while the first bracketed term in expression (15) is unaffected by the distribution of the outside value. Hence, the threshold value for which pooling is preferred to separating is smaller for $F_{\alpha'}$ than for $F_{\alpha''}$. ■

[28] This is the LP's payoff without investment distortions. For a sufficiently high second-period duration δ, the limited partner's payoff under a first-period separating contract would be strictly lower.

References

Axelson, U., P. Strömberg, and M. Weisbach (2009): "Why Are Buyouts Levered? The Financial Structure of Private Equity Funds," *Journal of Finance*, 64(4), 1549–1582.

Breslow, S., and P. Schwartz (2015): *Private equity funds: formation and operation*. Practicing Law Institute.

Casamatta, C. (2003): "Financing and Advising: Optimal Financial Contracts with Venture Capitalists," *Journal of Finance*, 58(5), 2059–2085.

Cornelli, F., and O. Yosha (2003): "Stage Financing and the Role of Convertible Securities," *Review of Economic Studies*, 70(1), 1–32.

Dewatripont, M., I. Jewitt, and J. Tirole (1999): "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies*, 66, 183–198.

Freixas, X., R. Guesnerie, and J. Tirole (1985): "Planning under Incomplete Information and the Ratchet Effect," *Review of Economic Studies*, 52, 173–192.

Gibbons, R., and K. Murphy (1992): "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, 100(3), 468–505.

Gompers, P. (1995): "Optimal Investment, Monitoring, and the Staging of Venture Capital," *Journal of Finance*, 50(5), 1461–1490.

Gompers, P., and J. Lerner (1999a): "An Analysis of Compensation in the U.S. Venture Capital Partnership," *Journal of Financial Economics*, 51, 3–44.

——— (1999b): *The Venture Capital Cycle*. MIT Press, Cambridge MA.

Harris, L. (2010): "A critical theory of private equity," *Delaware Journal of Corporate Law (DJCL)*, 35(1).

Hellmann, T. (1998): "The Allocation of Control Rights in Venture Capital Contracts," *RAND Journal of Economics*, 29(1), 57–76.

Holmström, B. (1999): "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66(1), 169–182.

Kaplan, S. N., and P. Strömberg (2003): "Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts," *Review of Economic Studies*, 70(2), 281–315.

Kaplan, S. N., and P. Strömberg (2004): "Characteristics, Contracts, and Actions: Evidence From Venture Capitalist Analyses," *Journal of Finance*, 59(5), 2173–2206.

Laffont, J., and J. Tirole (1988): "The Dynamics of Incentive Contracts," *Econometrica*, 56(5), 1153–75.

Laffont, J., and J. Tirole (1990): "Adverse Selection and Renegotiation in Procurement," *Review of Economic Studies*, 57, 597–625.

Lambert, R. (1993): "Long-Term Contracts and Moral Hazard," *Bell Journal of Economics*, 14(2), 441–452.

Legath, J. (2011): *Transaction and Monitoring Fees: On the Rebound?* Dechert and Preqin.

Lerner, J., F. Hardymon, and A. Leamon (2005): *Venture Capital and Private Equity: A Casebook*. John Wiley and Sons, Inc.

Lerner, J., and A. Schoar (2004): "The Illiquidity Puzzle: Theory and Evidence from Private Equity," *Journal of Financial Economics*, 72(1), 3–40.

Litvak, K. (2004): "Governing by Exit: Default Penalties and Walkaway Options in Venture Capital Partnership Agreements," *Willamette Law Review*, 40, 771–812.

——— (2009): "Venture Capital Partnership Agreements: Understanding Compensation Arrangements," *University of Chicago Law Review*, 76, 161–218.

Malcomson, J. M., and F. Spinnewyn (1998): "The Multiperiod Principal-Agent Problem," *Review of Economic Studies*, 55, 391–408.

Metrick, A., and A. Yasuda (2010a): "The economics of private equity funds," *Review of Financial Studies*, 23(6), 2303–2341.

——— (2010b): "The Economics of Private Equity Funds," *Review of Financial Studies*, 23(6), 2303–2341.

Monk, A. H., and R. Sharma (2015): "Re-Intermediating Investment Management: A Relational Contracting Approach," *Available at SSRN 2625303*.

Phalippou, L., C. Rauch, and M. P. Umber (2018): "Private Equity Portfolio Company Fees," *Journal of Financial Economics, Forthcoming*.

Rey, P., and B. Salanie (1990): "Long-Term, Short-Term and Renegotiation: on the Value of Commitment in Contracting," *Econometrica*, 58(3), 597–619.

Sahlman, W. (1990): "The Structure and Governance of Venture-Capital Organizations," *Journal of Financial Economics*, 27, 473–521.

Schmidt, K. M. (2003): "Convertible Securities and Venture Capital Finance," *Journal of Finance*, LVIII(3), 1139–1166.

Stone, H. (2009): *Raising Capital for Private Equity Funds: Leading Lawyers on Navigating Recent Fund Formation Trends, Understanding Legal and Contractual Hurdles*. Aspatore Books.

Toll, D. M., and V. Vayner (2012): *PE/VC Partnership Agreements Study 2012-2013*. Thomson Reuters Deals Group.