Default Penalties 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

June 30, 2016

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 show that default penalties are part of an optimal contract between limited and general partners. Default penalties help limited partners in screening general partners, and in minimizing distortions in investment levels and fees, caused by information asymmetries between general and limited partners. We also show that an optimal fee structure requires management fees that are proportional to capital under management, and transaction fees that are paid during the life of the fund when investments are made.

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

[1] We are especially grateful to Ulf Axelson, Catherine Casamatta, Xavier Freixas, Javier Gil-Bazo, Alexander Guembel, Demian Macedo, David Martinez-Miera, Alan Morrison, Ludovic Phalippou, Pablo Ruiz-Verdú, 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), and the Corporate Finance Day 2012 (Ghent). Sergio Vicente acknowledges support from the Spanish Ministry of Education and Culture, project ECO2013-42849-P, and 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*. That is, LPs do not hand over the entire committed capital at the outset but supply it in stages, retaining the real option to default on their commitment obligations. An LP may be forced or choose to default on a capital call for a number of reasons; for example because of 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 must specify a *default penalty* for the LP's failure to honor a capital call. Penalty clauses are often written as long lists of punishments, ranging from relatively mild to very severe, implying for instance a 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 venture (73%) and buyout funds (72%) include the most severe default penalty surveyed: 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)), the default provisions were subject

[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 in 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 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))

of little discussion in the past because LPs were assumed not to default on their obligations. Nowadays, however, and especially since the financial crisis, LPs are forced to consider their options when capital is called, including the decision to default on a capital call. Surprisingly, 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 (2004)). To our knowledge, and despite of their pervasiveness and importance, the finance literature has not formally analyzed the role default penalties play in private equity, how harsh they should be, nor on which fund characteristics they should depend on.

This paper offers a formal analysis of the role that default penalties play in the design of private equity funds. We construct a two-period adverse selection model in which a relatively large LP and a relatively small GP contract on the capital that should be invested in each period, the fees, and the penalties for dishonoring a capital call.[4] The model has three key ingredients: (i) long-term relationship without commitment - LPs and GPs contract over two periods but renegotiation is possible in the interim; (ii) asymmetric information - GP's costs to generate the investment opportunities and run the fund profitably are initially not known to the LP; and (iii) an outside option - an alternative investment opportunity may become available to the LP in the second period. We show that with optimal contracts better GPs run funds with larger capital investments, higher fees, and higher associated penalties, but are also expected to deliver greater profits. Default penalties are an integral part of this optimal contract because they improve the efficiency of investment decisions.

When a GP's ability is difficult to assess ex-ante, LPs need to design contracts so as to screen out better GPs from worse GPs. Better GPs end up running larger funds and producing larger profits in exchange for larger fees. Better GPs command larger fees not only because they manage larger pools of assets but also because they need to be provided with incentives to reveal their own ability. Indeed, better GPs could pretend to be worse GPs and claim to have higher costs to generate the same investment opportunities. As in the

[4] We deliberately abstract from moral hazard issues and focus on how adverse selection problems affect the choice of management and deal fees.

real world, GPs are compensated partly with fees that are proportional to the capital under management (management fees) and partly with fees that are non-proportional (such as the transaction fees). We show that the proportional fees should represent a smaller percentage of capital under management for larger funds. This is consistent with the finding of Gompers and Lerner (1999a) that management fees decrease with fund size. To induce better GPs to run larger funds with lower proportional fees, LPs offer them larger non-proportional fees, which represent the transaction fees in our model.

In a dynamic setting, initial screening is even more complicated than in a static setting. Separation between GPs can be achieved only by promising very high non-proportional fees to better GPs upfront. As shown by Laffont and Tirole (1988), in the absence of commitment, better GPs anticipate that information provided in the first period can be used to reduce informational rents in the second period. In the absence of default penalties, contracts need to include significant distortions. High up-front payments may give incentives to better GPs to separate out from worse GPs, but they also give incentives to worse GPs to pretend to be better GPs, thus creating "countervailing incentives." Indeed, 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 would need to over-invest in better GPs and pay them excessively high fees, and under-invest and under-pay worse GPs.

We show that the presence of alternative options reduces countervailing incentives. Better GPs can be paid less in terms of informational rents in the first period because LPs may not be willing to continue onto the second period. This reduces the incentives of worse GPs to pretend to be better GPs. In turn, distortions can be reduced but may not necessarily be eliminated. Notice that in a framework with zero outside options, as in Laffont and Tirole (1988), in equilibrium LPs will never use the zero default penalty to reduce GPs' fees, even if they could do so, because contracts are renegotiation-proof. But, in our framework, LPs use the zero default penalties to leave the fund in the interim if the alternative option is sufficiently attractive.

Default penalties can reduce investment distortions even further because they allow LPs to stagger the payments of non-proportional fees. Better GPs get the full payment of the informational rents in the second period, either in terms of transaction fees for running the fund for the entire duration of the partnership, or in terms of a default penalty in case the LP prefers to exit. Thanks to the default penalties, it is no longer necessary to pay the full rents upfront, as exit is not a threat to the rents of better GPs. Worse GPs have less incentives to pretend to be better GPs, as they would obtain an (underserved) reward only in the case of the LP exiting. In the case of continuation, worse GPs would not be able to obtain the full informational rents as they would refuse to run the fund in the second period. Therefore, unlike what happens with better GPs, worse GPs do not get complete insurance against the risk of continuation.

The model draws predictions on the use and the size of default penalties. The model relates default penalties to fund performance and size, to the degree of asymmetric information, to the probability of default, and to the residual life of the fund at the time of default. Consistent with the predictions of the 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.

To our knowledge, our paper provides a first rationale for the use of transaction fees. Most prior work on contractual terms of private equity funds has instead focused on management and performance-based fees. Management fees, typically representing 2% of committed capital, are supposed to cover the costs of managing the fund and generating investment opportunities. Performance-based fees, such as the carried interest of 20% of the profits, are meant to reduce the moral hazard problems arising from fund mismanagement by the GP (Lerner and Schoar (2004)). Significantly less understood are the transaction fees, which are charged by GPs to portfolio firms and are rarely fully rebated against management fees

(Phalippou, Rauch, and Umber (2015)).[5] 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." As for default penalties, being relatively unknown does not mean that they are not significant (Phalippou, Rauch, and Umber (2015)).[6]

The optimal contract in our setup includes proportional and non-proportional fees, which we interpret as management and transaction fees, respectively. Carried interest could be incorporated in the model if one extended it to account for GPs' moral hazard. The model shows that management fees should represent a smaller percentage in larger funds. Transaction fees, instead, should be bigger for better GPs and should be paid not at the inception of the fund, but later on. Phalippou, Rauch, and Umber (2015) and Legath (2011) highlight the importance of transaction fees in the LBO industry. They show that transaction fees represent an important source of revenues for GPs because they are computed as a non-trivial percentage of the size of the deals.

The literature on private equity has dedicated a great deal of attention to the relationship between venture capitalists and entrepreneurs (Casamatta (2003), Cornelli and Yosha (2003), Gompers (1995), Hellmann (1998), Kaplan and Strömberg (2003), Kaplan and Strömberg (2004), and Schmidt (2003)). However, there has been little research on the design of partnership agreements.[7] Axelson, Strömberg, and Weisbach (2009), which we refer to as ASW hereafter, is perhaps the closest to our line of investigation. 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 his incentives to invest in bad deals, since bad deals contaminate his stake in the good deals. Instead, we focus our model on the problem of screening GPs with heterogeneous ability, and show how

[5] In our definition, transaction fees also include consulting, advisory and other related fees charged to portfolio companies.

[6] 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.

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

the screening problem is affected by the dynamic nature of the relationship between GPs and LPs. In the same spirit, the focus of our model is more on transaction fees than on carried interest, as instead is the case of Lerner and Schoar (2004).

In terms of theory, this paper builds on Laffont and Tirole (1988), who study a two-period adverse selection problem with no commitment.[8] As in their set-up, the agent can (although in equilibrium will *not*) unilaterally interrupt the partnership in the interim, at no cost. In our framework, however, the principal may have to pay a penalty to unilaterally interrupt the partnership, and therefore the degree of commitment between the principal and the agent is asymmetric. Moreover, the default penalty is an integral part of the contract, and therefore the degree of commitment is endogenous. Of course, paying the default penalty and exiting will only make sense in practice if there is a positive outside option.[9]

The rest of the paper is organized as follows. We describe the model in section 2. Section 3 lays out a description of the benchmark contracts. We devote section 4 to an evaluation of the distortions induced by countervailing incentives in the absence of penalties. 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 it to current available evidence and suggest potential avenues for research. Section 8 concludes.

2 The model

Consider a two-period environment in which there is an LP (the Principal), who has capital to invest, and a penniless GP (the Agent), who has investment opportunities. The LP and the GP may engage in a private equity partnership, also referred to as *the fund*, which is managed

[8] Indirectly, we borrow ideas about dynamic contracts from Dewatripont, Jewitt, and Tirole (1999), Holmström (1999), Gibbons and Murphy (1992), Lambert (1993), Malcomson and Spinnewyn (1998) and Rey and Salanie (1990). See also Bernardo, Cai, and Luo (2004) for an application of commitment to capital budgeting.

[9] Our set-up also differs from Laffont and Tirole (1990) because they assume that the contracting parties cannot unilaterally leave the partnership in the interim. That is, they effectively assume that both the LP and the GP would have to pay an "infinite" penalty to unilaterally leave the partnership.

by the GP. The LP must decide which amount k_t to contribute to the fund in each period $t = \{1, 2\}$. In each period, a capital investment k generates a net return of $R(k)$, where R is a publicly observable function, with $R(0) = 0$, $R' > 0$ and $R'' < 0$. We also assume that R satisfies the Inada conditions $R'(0) = +\infty$ and $\lim_{k\to 0} R'(k) \cdot k = 0$.[10] Generating $R(k)$ for a fund of size k entails a cost $\theta_i \cdot k$ to the GP, where $\theta_i \in \{\theta_g, \theta_b\}$, with $\Delta\theta \equiv \theta_b - \theta_g > 0$, a set-up that reflects that *good* GPs can get the same investment opportunities as *bad* GPs at a lower cost.[11] The type of the GP is his own private information, but it is common knowledge that it is $i \in \{g, b\}$ with probability υ_i, with $\upsilon_g + \upsilon_b = 1$. Second-period payoffs are "discounted" by a factor $\delta \geq 0$ which, as it is standard, may be larger than 1 to represent cases where the second period lasts much longer than the first period. 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 LP's outside option is 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 a net present value of I is missed. The value of I is the realization of a random variable distributed with a common knowledge cumulative distribution function F, with an expected value $\mu \equiv \displaystyle\int_{\underline{I}}^{\overline{I}} I \cdot dF(I)$. The realization of I takes place at the beginning of the second period and is not contractible upon, either because it is not publicly observable or because enforcing it in a court of law is prohibitively costly. Since we focus on the impact of the changing investment opportunities on the side of the LP we assume, as it is standard in the literature, that the GP's outside option is zero in both periods.

We assume that partnership agreements are designed by the LP and come as a take-it-or-leave-it offer to the GP.[12] A contract $C_i = \{k_{1i}, k_{2i}, x_{1i}, x_{2i}, P_i\}$ designed for a type$-i$ GP

[10] As we shall see below, these conditions ensure that a positive investment level is always optimal.

[11] Monk and Sharma (2015), for example, state that (page 10) "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." Management fees in our model (which we interpret as the part of the fees that is proportional to the size of the fund) are also such that they cover the costs of running the fund. In our model, the proportional fees should in fact represent a smaller percentage of capital under management for larger funds, which is also consistent with the finding of Gompers and Lerner (1999a) that management fees decrease with fund size.

[12] 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

specifies, for each round t, where $t = 1, 2$, the LP's capital contribution $k_{ti} \geq 0$ to the fund, a fee $x_{ti} \geq 0$ to be paid to the GP and a penalty $P_i \geq 0$ that the LP must pay to the GP if he decides to interrupt the partnership before the second period.[13] This specification potentially allows for pooling contracts, that is, menus of contracts such that $C_g = C_b$. Throughout the text, we call *optimal* contracts those that maximize the LP's profits. As we show below, fees x_{ti} in optimal contracts have a two-part structure. One part, which we shall interpret as *management fees*, is proportional to the size of the fund that the GP runs. The second part, which we shall interpret as *transaction fees*, is non-proportional to the fund size.

The LP's profit from signing a contract $C = \{k_1, k_2, x_1, x_2, P\}$ is given by:

$$\Pi_L(C) = \begin{cases} \sum\limits_{t=1,2} \delta_t \cdot (R(k_t) - x_t) & \text{if the partnership extends to } t = 2 \\ R(k_1) - x_1 + \delta \cdot (I - P) & \text{if the partnership is broken before } t = 2 \end{cases} .$$

Similarly, a type$-i$ GP's payoff, for $i \in \{g, b\}$, is given by:

$$\Pi_i(C) = \begin{cases} \sum\limits_{t=1,2} \delta_t \cdot (x_t - \theta_i \cdot k_t) & \text{if the partnership extends to } t = 2 \\ x_1 - \theta_i \cdot k_1 + \delta \cdot P & \text{if the partnership is broken before } t = 2 \end{cases} .$$

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.

[13] When choosing to default on an obligation to a fund, an LP must also consider the effects on his own reputation. For simplicity, we design the optimal contract abstracting from reputational considerations, as 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 the full range of alternative private equity investment options.

The surplus created by a partnership between an LP and a type$-i$ GP is therefore given by:

$$\pi(C)=\begin{cases}\sum_{t=1,2}\delta_t\cdot(R(k_t)-\theta_i\cdot k_t) & \text{if the partnership extends to } t=2\\ R(k_1)-\theta_i\cdot k_1+\delta\cdot I. & \text{if the partnership is broken before } t=2\end{cases}.$$

Notice that both x_t and P are mere transfers of resources from the LP to the GP, so that the surplus generated by the partnership does not depend on either of them.

We assume that LPs and GPs can contract over two periods but that there is no commitment.[14] LPs, instead, 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 mainly in providing capital to invest. 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.

The timing of contracting is as follows. At the beginning of period $t=1$, the LP offers a menu of contracts $C=\{C_g,C_b\}$, potentially allowing for pooling contracts (i.e., $C_g=C_b$). Upon acceptance, a type$-i$ GP manages a fund of size k_{1i} in exchange for a fee x_{1i}. At the end of period $t=1$, the first round ends and payoffs, as specified above, are realized. At the beginning of period $t=2$, either party may opt out of the partnership but, if specified in the contract, the LP must pay a penalty to the GP to do so. In the event of the partnership being interrupted, parties can either get their respective outside option at $t=2$ or sign a new contract for the remaining period. Indeed, the contracting parties can renegotiate the contribution to the fund and the fees to k'_{2i} and x'_{2i}, respectively. If the partnership is not interrupted, the LP contributes an amount k_{2i} to the fund in exchange for a transfer of x_{2i} as specified in the original contract. Upon acceptance of either the original or the new 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.

[14]It seems hard for an LP to force a GP to run a fund, even if contractually obliged to do so, when the GP could be better off outside the partnership. The mere threat of mismanaging the fund should suffice to convince the LP to let the GP opt out of the fund. However, as we shall see below, in equilibrium GPs never opt out of the partnership, even though they are allowed to do so.

Notice that, as in Laffont and Tirole (1988), contracts must be *renegotiation-proof*, as any contract subject to a Pareto improvement can be freely terminated and renegotiated in the interim.

3 Benchmark contracts

Before proceeding with the analysis of the problem at hand, we first lay out benchmark contracts, which help dissect the distortions induced by asymmetric information, as well as to introduce some notation that will be helpful for the following sections. First, we describe the *first-best contract*, in which the LP is perfectly informed of the GP's type. Second, we provide the well-known characterization of the *second-best static contracts (S)*, which serve as the benchmark to construct dynamic contracts. Third, we characterize the *full-commitment contracts (FC)*, in which the LP and the GP commit to abide by the terms of a two-period contract even if there is room for improvement at some interim period at which some uncertainty is resolved. As noted by Dewatripont (1988), the full-commitment contract is not renegotiation-proof, so that partners will be willing to renegotiate its terms if it is mutually beneficial to do so. However, it serves as a benchmark to understand which would be the profit-maximizing investment levels within the partnership in a two-period setting.[15] As these contracts are fairly standard, here we simply state the main results and relegate the formal derivation to the appendix. Further details can be found in Laffont and Martimort (2002).

[15] In a sequel of seminal papers, Laffont and Tirole (1987), Laffont and Tirole (1988) and Laffont and Tirole (1990) show that the full-commitment contract maximizes the principal's payoff in a setting that differs from ours mainly in that $I = 0$. The main insight from these papers is that full commitment contracts maximize the LP's payoff. Although renegotiating the contract after types revelation occurs in the first period would increase efficiency ex-post, the renegotiation prospects would harden incentives ex-ante, thereby reducing the LP's payoff overall.

3.1 First-best contracts

Suppose first that the GP's type is known to the LP. First-best investment levels k_i^* satisfy $R'(k_i^*) = \theta_i$ in both $t = 1, 2$, which are the investment levels that equate the marginal cost to the marginal return. The surplus generated by an efficient static contract is given by

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

As shown in the appendix, the good GPs runs a larger fund (i.e., $k_g^* > k_b^* > 0$) and generates a larger surplus (i.e., $\pi_g^* > \pi_b^* > 0$). With full information, two-period contracts consist of a sequence of first-best static contracts. A partnership with a type$-i$ GP is terminated before the second period if and only if $I > \pi_i^*$, that is, when the outside option is worth more than the surplus that can be generated within the partnership.

3.2 Second-best static contracts

Let us from now on consider the asymmetric information case. In a static framework, the LP solves the following optimization program:

$$\left.\begin{array}{lll} \max_{\{k_i \geq 0, x_i \geq 0\}_{i\in\{g,b\}}} & \sum\limits_{i=g,b} \upsilon_i \cdot \left[\left(R\left(k_i\right) - x_i\right)\right] & \\ s.t & x_i - \theta_i \cdot k_i \geq 0 & [PC.i] \\ & x_i - \theta_i \cdot k_i \geq x_j - \theta_i \cdot k_j & [ICC.i] \end{array}\right\}, \tag{S}$$

where $(PC.i)$ and $(ICC.i)$ stand for type$-i$'s participation and incentive compatibility constraints, respectively.

Whether this program is solved by a pooling or by a separating contract depends on the ex-ante likelihood that the GP is good. Let us assume from now on that separating contracts dominate pooling contracts in a static context.[16]

Assumption 1 (Proportion of Good GPs) $\nu_g \geq \frac{\pi_b^* \cdot}{\pi_g^*}$.

As in the case with perfect information, the LP would like to assign a first-best level k_i^* of investment to both types of GPs and compensate each of them with the level of fees $\theta_i \cdot k_i^*$ so that they can break even. However, if this menu of contracts were offered, the good GP would have an incentive to claim to be a bad GP, so as to get a net payoff $\Delta\theta \cdot k_b^*$, because she would be able to generate the same investment opportunities at a lower cost. To reduce the rent paid to the good GP, the optimal menu of contracts requires a downward distortion of the optimal investment for the bad GP. The profit-maximizing contract addresses the ensuing trade-off between efficiency and rents.

In equilibrium, the good type's incentive compatibility and the bad type's participation constraints are the only ones that are binding. As a result, the program simplifies to:

$$\max_{\{k_i \geq 0\}_{i=g,b}} \upsilon_g \cdot (R(k_g) - \theta_g \cdot k_g - \Delta\theta \cdot k_b) + \upsilon_b \cdot ([(R(k_b) - \theta_b \cdot k_b)]) \tag{S'}$$

and thus investments are $k_g^S = k_g^*$ and $k_b^S < k_b^*$, which is characterized by $R'(k_b^S) = \theta_b + \frac{\nu_g}{\nu_b} \cdot \Delta\theta$, and associated fees $x_g^S = \theta_g \cdot k_g^* + \Delta\theta \cdot k_b^S$ and $x_b^S = \theta_b \cdot k_b^S$. The LP's profit is then given by:

$$\Pi_L^S \equiv \upsilon_g \cdot \left(\pi_g^* - \Delta\theta \cdot k_b^S\right) + \upsilon_b \cdot \pi_b^S. \tag{2}$$

A bad GP does not have incentives to claim to be a good GP, as the gains $\Delta\theta \cdot k_b^S$ in terms of fees would be outweighed by the cost $\Delta\theta \cdot k_g^*$ of running the fund designed for the good GP.

3.3 Full-commitment two-period contracts

In this section, we characterize the *full-commitment menu of contracts* $C^{FC} = \left\{C_g^{FC}, C_b^{FC}\right\}$. Crucially, the full-commitment contract neither allows parties to renegotiate the terms of the

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

contract, nor allows them to opt out of the partnership with or without paying a default penalty.

The LP solves the following problem:

$$\left.\begin{array}{lll} \max_{\{k_{1i}\geq 0, x_{1i}\geq 0\}_{\substack{t=1,2\\ i=g,b}}} & \sum_{t=1,2}\delta_t\cdot\sum_{i=g,b}\upsilon_i\cdot\left[\left(R\left(k_{ti}\right)-x_{ti}\right)\right] & \\ s.t & \sum_{t=1,2}\delta_t\cdot\left(x_{ti}-\theta_i\cdot k_{ti}\right)\geq 0 & [PC.i] \\ & \sum_{t=1,2}\delta_t\cdot\left(x_{ti}-\theta_i\cdot k_{ti}\right)\geq\sum_{t=1,2}\delta_t\cdot\left(x_{tj}-\theta_i\cdot k_{tj}\right) & [ICC.i] \end{array}\right\} \tag{FC}$$

As shown in Laffont and Tirole (1988), the optimal contract consists of a repetition of the static second-best contracts discussed above, that is: $k_{tg}^{FC}=k_g^*$ and $k_{tb}^{FC}=k_b^S$ for $t=1,2$ and intertemporal fees given by $x_{1g}^{FC}+x_{1g}^{FC}=\sum_{t=1,2}\delta_t\cdot\left(\theta_g\cdot k_{tg}+\Delta\theta\cdot k_{tb}\right)$ and $x_{1b}^{FC}+x_{1b}^{FC}=\sum_{t=1,2}\delta_t\cdot\theta_b\cdot k_{tb}$. Observe that the contract cannot pin down each period's transfer, but simply defines an intertemporal transfer for each type of GP. We summarize this discussion in the following proposition.

Proposition 1 (Full-commitment contracts) *The two-period full-commitment menu of contracts consists of a repetition of the static second-best contracts.*

As shown in Laffont and Tirole (1988), two-period full-commitment contracts have the virtue that the most profitable static arrangement for the LP can be replicated for two periods. But in our setting, the contract cannot be made contingent on the realization of I and it may therefore impose a high ex-post loss in the case that the LP's second-period outside investment opportunity happens to be large. As a result, full-commitment contracts do not always dominate the no-commitment contracts, to which we turn next.[17]

[17] Naturally, there are two other ways to understand full commitment. One way would be committing to a one-period contract only, agreeing that, no matter how low the outside option for the LP may be in the second period, continuation will not occur. Another one would be offering a menu of separating contracts, one lasting for just one period, the other one lasting for two periods, so that continuation would occur with only one type of agent. Full commitment, intended in the traditional sense as a two-period non-renegotiable contract, seems more in line with the way it has been addressed in the literature. In any case, these other contracts are also dominated by no-commitment contracts for some distributions of the outside option.

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 *separating contracts with zero default penalties* $C^Z = \left\{C_g^Z, C_b^Z\right\}$. In the following section we compare them with those obtained with optimally-set default penalties.[18]

4.1 Contract design

We first determine the contractual terms of the second period. Since in separating contracts types are revealed after the first period, contracts need to prescribe 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: $k_{2i}^Z = k_i^*$, $i = g, b$.[19] Since all private information is revealed in the first period, there is no informational rent in the second period, i.e., $x_{2i}^Z = \theta_i \cdot k_i^*$. Hence, the GP makes zero profits in the second period, as the LP extracts the entire surplus π_i^* from the partnership. Since the relationship may unilaterally be broken by the LP before the second period at no cost, she would interrupt the partnership whenever $I > \pi_i^*$.

The consequences in terms of incentives are the following. First notice that good GPs could obtain a discounted (net) payoff of $\delta \cdot \Delta\theta \cdot k_b^*$ if they were to pretend to be a bad GP *and* the partnership was extended into the second period, an event which would occur with probability $F(\pi_b^*)$. On the other hand, as in the case of static second-best contracts, a bad type would never be willing to run a fund of size k_g^* in the second period in exchange for a transfer of $\theta_g \cdot k_g^*$ as he would incur a net loss of $\Delta\theta \cdot k_g^*$. Hence, an impersonation of a good GP by a bad one would only last for the first period. Hence, the LP would solve the following

[18] By separating contracts we refer to contracts in which each GP type takes a different contract, hence eliciting his 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. As we shall see in section 6, the optimal default penalty in pooling contracts is zero.

[19] Any contract in which separation of types is achieved in the first period would yield the efficient outcomes in the second period. This phenomenon, first coined by Freixas, Guesnerie, and Tirole (1985) as the *racthet effect*, was identified in this context by Laffont and Tirole (1987) and Laffont and Tirole (1988).

problem:

$$\left.\begin{array}{lll}\max_{\{k_{1i}\geq 0,x_{1i}\geq 0\},i=g,b} & \sum_{i=g,b}\upsilon_i\cdot\left(R\left(k_{1i}\right)-x_{1i}\right)+\delta\cdot\Pi_{2L}^{Z} & \\ s.t & x_{1i}-\theta_i\cdot k_{1i}\geq 0 & [PC.i] \\ & x_{1g}-\theta_g\cdot k_{1g}\geq x_{1b}-\theta_g\cdot k_{1b}+\delta\cdot\Delta\theta\cdot F\left(\pi_b^*\right)\cdot k_b^* & [ICC.g] \\ & x_{1b}-\theta_b\cdot k_{1b}\geq x_{1g}-\theta_b\cdot k_{1g} & [ICC.b]\end{array}\right\}, \quad \text{(Z)}$$

where

$$\Pi_{2L}^{Z}\equiv\sum_{i=g,b}\upsilon_i\cdot\left(F\left(\pi_i^*\right)\cdot\pi_i^*+\int_{\pi_i^*}^{\bar{I}}I\cdot dF\left(I\right)\right)$$

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

The key aspect to recognize here is that, since all private information is revealed in the first period and the LP can opt out of partnership at no cost before the second period, the LP can extract the entire surplus from the GP if the partnership lasts for two periods. Therefore, any incentive compatible contract will have to include an informational rent *upfront*, and equal to $\Delta\theta\cdot\left(k_{1b}+\delta\cdot F\left(\pi_b^*\right)\cdot k_b^*\right)$, to a good GP for him to be willing to elicit his true type. In what follows, we show that first-period investments may have to be distorted away from the static second-best investment levels. Moreover, we identify the role of the LP's outside option in the conditions under which this may occur.

4.2 No-distortions separating contract

Suppose first that the bad type's incentive compatibility constraint does not bind, as in the static problem. Since the good type's incentive compatibility constraint and the bad type's participation constraint bind, as usual, the LP's problem turns into the following:

$$\max_{\{k_{1i}\}_{i=g,b}}\upsilon_g\cdot\left(R\left(k_{1g}\right)-\theta_g\cdot k_{1g}-\Delta\theta\cdot\left(k_{1b}+\delta\cdot F\left(\pi_b^*\right)\cdot k_b^*\right)\right)+\upsilon_b\cdot\left(\left[\left(R\left(k_{1b}\right)-\theta_b\cdot k_{1b}\right)\right]\right)+\delta\cdot\Pi_{2L}^{Z}$$

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

Definition 1 (No-distortions Separating Contracts) *The* no-distortions separating contracts $C^{ND} = \left\{C_g^{ND}, C_b^{ND}\right\}$ *consist of the second-best static investments in the first period (i.e., $k_{1g}^{ND} = k_g^*$ and $k_{1b}^{ND} = k_b^S$), efficient investments in the second period (i.e., $k_{2g}^{ND} = k_g^*$ and $k_{2b}^{ND} = k_b^*$) and fees $x_{1g}^{ND} = \theta_g \cdot k_g^* + \Delta\theta \cdot \left(k_b^S + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right)$ and $x_{1b}^{ND} = \theta_b \cdot k_b^S$ in the first period and $x_{2i}^{ND} = \theta_i \cdot k_i^*$ in the second.*

However, the bad type's incentive compatibility constraint may be binding. Indeed, if offered the no-distortions separating contracts C^{ND}, a bad GP would obtain a zero payoff by taking his own contract, as his participation constraint would be binding. But he could pretend he is a good GP, run a fund of size k_g^* in the first period, refuse to run the fund in the second period, and obtain a net payoff of $\Pi_b\left(C_g^{ND}\right) \equiv \Delta\theta \cdot \left(k_b^S + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^* - k_g^*\right)$. Hence, if $\Pi_b\left(C_g^{ND}\right) > 0$, C^{ND} is not incentive-compatible, as a bad type would impersonate a good type. Unlike in static contracts, the high compensation package that a good GP must be given to take his own contract might induce a bad GP to pretend he is good. This observation motivates the following definition.

Definition 2 (Countervailing Incentives) *There are* countervailing incentives *when the* no-distortions separating contract C^{ND} *is not incentive-compatible.*

In the following lemma we characterize an *incentive-compatibility condition* for any contract with zero penalties and relate it to countervailing incentives.

Lemma 1 (Incentive-Compatibility Condition) *A pair of investment levels (k_{1b}, k_{1g}) involving separation of types in the first period is incentive compatible if and only if:*

$$F\left(\pi_b^*\right) \leq \frac{k_{1g} - k_{1b}}{\delta \cdot k_b^*}. \tag{3}$$

This expression holds as equality if and only if there are countervailing incentives.

Proof. See Appendix. ■

This condition prescribes that investment levels for each type must be sufficiently separated away from each other, so that it is prohibitively costly for a bad type to run a large size fund. The following corollary is a straightforward implication of Lemma 1.

Corollary 1 (Range of No-Distortions with Zero Default Penalties) *The no-distortions separating contracts menu C^{ND} is incentive-compatible if and only if:*

$$F\left(\pi_b^*\right) \leq F_b^{CI} \equiv \frac{k_g^* - k_b^S}{\delta \cdot k_b^*}. \tag{4}$$

The threshold F_b^{CI} establishes a maximum value for the probability $F\left(\pi_b^*\right)$ so that there are no countervailing incentives. When condition (4) holds, the probability of continuation with a bad GP is small enough so that the informational rent that needs to be paid to the good GP to induce separation is not attractive enough for the bad GP. In this case, the no-distortions separating menu of contracts is optimal.

Observe that we can fix $F\left(\pi_b^*\right)$ and write Corollary 1 in terms of second-period duration. For any given $F\left(\pi_b^*\right) > 0$ the no-distortions separating contracts menu C^{ND} is optimal if and only if $\delta > \delta^{CI} \equiv \frac{k_g^* - k_b^S}{F\left(\pi_b^*\right) \cdot k_b^*}$, that is, the second period lasts for a sufficiently long period of time. Although we will perform most comparisons relying on F_b^{CI}, we will refer to δ^{CI} when comparing pooling and separating contracts.

In Laffont and Tirole (1988), the absence of second-period profitable outside options makes the probability of continuation onto the second period equal to 1. Hence, in their setting, there are countervailing incentives if and only if $F_b^{CI} < 1$, i.e., the cost of pretending to be a good GP $\Delta\theta \cdot k_g^*$ is lower than the benefits in terms of informational rents $\Delta\theta \cdot (k_b^S + \delta \cdot k_b^*)$. In our set-up, there is an additional condition to have countervailing incentives, namely that $F_b^{CI} < F\left(\pi_b^*\right)$, since the informational rent is strictly smaller. Therefore, the no-distortion contracts are easier to implement.

4.3 Investment distortions

Consider now the case in which there are countervailing incentives, i.e., condition (4) is not satisfied. Then, it follows from Lemma 1 that the bad GP's incentive compatibility constraint binds, so that $k_{1g}^{Z} = k_{1b} + \delta \cdot F(\pi_b^*) \cdot k_b^*$. Substituting into the good GP's incentive compatibility constraint, we have that $x_{1g}^{Z} = \theta_b \cdot k_{1g}^{Z}$. Therefore, we can write the LP's problem as:

$$\max_{k_{1b} \geq 0} \upsilon_g \cdot \left(R\left(k_{1b} + \delta \cdot F(\pi_b^*) \cdot k_b^*\right) - \theta_b \cdot \left(k_{1b} + \delta \cdot F(\pi_b^*) \cdot k_b^*\right) + \upsilon_b \cdot \left(R(k_{1b}) - \theta_b \cdot k_{1b}\right) + \delta \cdot \Pi_{2L}^{Z}.$$

The optimal level of investment for the bad type k_{1b}^{Z} satisfies $\upsilon_g \cdot R'(k_{1b}^{Z} + \delta \cdot F(\pi_b^*) \cdot k_b^*) + \upsilon_b \cdot R'\left(k_{1b}^{Z}\right) = \theta_b$. Notice that the conditions for the optimal investment levels in the static separating contract are $R'\left(k_g^*\right) = \theta_g$ and $R'(k_b^S) = \theta_b + \frac{\nu_g}{\nu_b} \cdot \Delta\theta$ and therefore $\upsilon_g \cdot R'\left(k_g^*\right) + \upsilon_b \cdot R'\left(k_b^S\right) = \theta_b$. Hence, the optimal distortion verifies:

$$\upsilon_g \cdot R'\left(k_{1b}^{Z} + \delta \cdot F(\pi_b^*) \cdot k_b^*\right) + \upsilon_b \cdot R'\left(k_{1b}^{Z}\right) = \upsilon_g \cdot R'\left(k_g^*\right) + \upsilon_b \cdot R'\left(k_b^S\right), \tag{5}$$

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

Proposition 2 (Optimal Separating Contracts with Zero Default Penalties) *The optimal menu of separating contracts with zero default penalties is as follows:*

(i) If $F(\pi_b^*) \leq F_b^{CI}$ (no countervailing incentives), the no-distortions menu of contracts C^{ND} is optimal. In particular, first-period investments satisfy $k_{1g}^{Z} = k_g^*$ and $k_{1b}^{Z} = k_b^S$.

(ii) If $F(\pi_b^*) > F_b^{CI}$ (countervailing incentives), the no-distortions menu of contracts C^{ND} is not incentive-compatible. The optimal first-period investments k_{1g}^{Z} and k_{1b}^{Z} are (uniquely) determined by $k_{1g}^{Z} = k_{1b}^{Z} + \delta \cdot F(\pi_b^*) \cdot k_b^*$ and expression (5), which implies both an upward and

a downward distortion of the good and the bad GP's investment levels, respectively, that is:

$$k_{1b}^Z < k_b^S < k_g^* < k_{1g}^Z.$$

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

(iv) Fees are given by $x_{1g}^Z = \underbrace{\theta_g \cdot k_{1g}^Z}_{Management\ Fees} + \underbrace{\Delta\theta \cdot \left(k_{1b}^Z + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right)}_{Transaction\ Fees}$ *and* $x_{1b}^Z = \underbrace{\theta_b \cdot k_{1b}^Z}_{Management\ Fees}$ *in the first period and* $x_{2i}^Z = \underbrace{\theta_i \cdot k_i^*}_{Management\ Fees}$ *in the second.*

(v) Consequently, the LP extracts all the surplus from a bad GP and pays an informational rent $\Pi_g^Z \equiv \Delta\theta \cdot \left(k_{1b}^Z + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right)$ *to a good GP.*

The following corollary is immediate:

Corollary 2 (Efficient Partnership Termination with Zero Default Penalties) *A partnership with a type–i GP will continue onto the second period if and only if the LP's outside investment opportunity generates a smaller surplus than the partnership with a type–i GP, that is, if and only if* $I \leq \pi_i^*$.

In the second period, the LP would renegotiate any contract so as to achieve the efficient investment levels and extract the entire surplus created by the partnership. Anticipating this, a good GP would require an up-front informational rent of Π_g^Z in order to elicit his true type. Hence, a good GP's contract would specify a first-period investment level k_{1g}^Z in exchange for an up-front payment of $\theta_g \cdot k_{1g}^Z + \Pi_g^Z$. By mimicking a good type, a bad GP would obtain a net payoff of $\Pi_g^Z - \Delta\theta \cdot k_{1g}^Z$. Notice that $\Delta\theta \cdot \left(k_{1b}^S + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right) - \Delta\theta \cdot k_g^* < 0$ if and only if $F\left(\pi_b^*\right) \leq F_b^{CI}$. Hence, when $F\left(\pi_b^*\right) > F_b^{CI}$ holds, the investment levels of the no-distortion contracts k_b^S and k_g^* are not implementable, as a bad GP would pretend he is a good one. In this case, k_{1b}^Z and k_{1g}^Z must be set so that $\Pi_g^Z - \Delta\theta \cdot k_{1g}^Z = 0$. Therefore, in order to induce a bad GP to truthfully reveal his type, a good GP's investment would have to be distorted upward above its efficient level, while a fund run by a bad GP would have to invest below the amount specified in the static second-best contract.

We can now compare the optimal contract here with that of Laffont and Tirole (1988), in which the outside opportunities are zero. Since the partnership may be interrupted in our set-up, the fees that must be paid to the good GP to induce separation are smaller, making the contracts of good GPs less attractive to bad GPs. As a result, on the extensive margin, distortions are less likely than in the absence of outside options.[20] On the intensive margin, the distortion is strictly smaller in our set-up than in the case in which the outside option is zero for certain. In our set-up, the distortion requires that $k_{1g}^Z - k_{1b}^Z = \delta \cdot F(\pi_b^*) \cdot k_b^*$, which is strictly smaller than the distortion required in the absence of outside options as long as $F(\pi_b^*) < 1$.

5 Separating contracts with optimal penalties

In what follows, we characterize the *separating contracts with optimally-set default penalties* $C^D = \left\{C_g^D, C_b^D\right\}$ and compare them to the contracts with zero default penalties analyzed in the previous section. We proceed as follows. We first analyze the size of the optimal penalties and the conditions under which partnerships are terminated after the first period. Then we address the conditions under which contracts with optimal penalties avoid first-period investment distortions. As we shall see, the use of penalties expands the range for which investments need not be distorted. Finally, we show that distortions, when needed, are always strictly smaller than in the case of zero penalties.

5.1 Default penalties

Let $\Pi_i^D \equiv \Pi_i\left(C_i^D\right)$ be the payoff, i.e., the informational rent, that a type$-i$ GP obtains in the case he takes his own contract. For convenience, we shall decompose it into first- Π_{1i}^D and

[20] In particular, whenever $F_b^{CI} < F(\pi_b^*) < 1$, contracts will have to be distorted in both set-ups. However, if $F(\pi_b^*) < F_b^{CI} < 1$, contracts will only introduce distortions in the case of no outside options, i.e. in Laffont and Tirole (1988). When $F_b^{CI} > 1$, the no-distortions menu of contracts is implementable in both settings.

second-period Π_{2i}^D informational rents, i.e.,:

$$\Pi_i^D \equiv \sum_{t=1,2} \delta_t \cdot \Pi_{ti}^D. \tag{6}$$

As in the case of zero penalties, the optimal contract entails efficient investment levels in the second period, as otherwise it would not be renegotiation-proof, that is, $k_{2i}^D = k_i^*$. Hence, a contract establishing a fee of x_{2i}^D would lead to a second-period informational rent of:

$$\Pi_{2i}^D \equiv \max\left\{0, x_{2i}^D - \theta_i \cdot k_i^*\right\} \tag{7}$$

for a type$-i$ GP. The value of 0 stands for the possibility that the GP does not run the fund in the second period, which would occur if running it yielded a negative payoff for him.

The following proposition establishes an important result.

Proposition 3 (Optimal Default Penalty) *In any incentive-compatible menu of contracts, the optimal default penalty P_i that the LP must pay to interrupt the partnership must be equal to the second-period informational rent, that is:*

$$P_i = \Pi_{2i}^D. \tag{8}$$

As a result, the GP would get the same second-period informational rent regardless of whether the LP opts out of the partnership or not.

The proof for 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 $x_{2i}^* = \theta_i \cdot k_i^*$, effectively reducing the GP's net rent in the second period 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 opt out of the partnership, which the GP would accept, since rejecting the offer would yield a (smaller) rent Π_{2i}^D to him.

The following corollary analyzes the conditions under which the LP would be willing to opt out of the partnership after the first period.

Corollary 3 (Efficient Partnership Termination with Default Penalties) *In a separating contract with default penalties, the LP opts out of partnerships efficiently, that is, whenever*

$$I > \pi_i^*. \tag{9}$$

Proof. See Appendix. ■

The LP's strategy entails opting out of the partnership whenever the realization of her outside option is larger than the surplus created by the second-period fund investment. Therefore, termination of partnerships are efficient, as in the case of zero penalties. The intuition for this result is as follows. In the case of zero penalties, interruptions are efficient because the LP extracts the entire second-period surplus, as all informational rents must be paid upfront. Hence, staying in the partnership yields a payoff of π_i^*, whereas opting out grants her a payoff of I. Hence, partnerships extend to the second period if and only if they generate a higher surplus than the LP's outside option, i.e., $I > \pi_i^*$. When the LP is required to pay a default penalty, staying in the partnership yields a payoff of $\pi_i^* - \Pi_{2i}^D$, whereas interrupting leads to a payoff of $I - P_i$. But, as we have shown in Proposition 3, the default penalty equals the second-period informational rent, that is, $P_i = \Pi_{2i}^D$. Therefore, when deciding whether to call for a partnership interruption, she simply compares the surplus created by the partnership with her outside option, i.e., $I > \pi_i^*$, as in the case without penalties.

5.2 Contract design

We now proceed to the analysis of the constraints of this program. Since a type$-i$ GP is fully insulated against any realization of the outside option, he would in any case obtain a net payoff of P_i in the second period if he takes a contract designed for himself. Hence, we can write a type$-i$ intertemporal participation constraint as:

$$x_{1i} - \theta_i \cdot k_{1i} + \delta \cdot P_i \geq 0. \tag{10}$$

In order to construct the 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$-j$ one. If the partnership is broken by the LP, then he gets P_j. But if, on the contrary, the LP decides not to opt out of the partnership, then a type$-i$ GP obtains a payoff of $x_{2j}^D - \theta_i \cdot k_j^*$, if he decides to stay in the partnership, or a payoff of 0, in case he opts out of the partnership after the first period. On the contrary, by taking a contract designed for himself, a type$-i$ GP gets fully insulated, as he receives the same payoff regardless of whether the partnership continues onto the second period. Type$-i$ incentive compatibility constraint then reads as follows:

$$\begin{aligned} &x_{1i} - \theta_i \cdot k_{1i} + \delta \cdot P_i \geq \\ &x_{1j} - \theta_i \cdot k_{1j} + \delta \cdot \left[F\left(\pi_j^*\right) \cdot \max\left\{x_{2j} - \theta_i \cdot k_j^*, 0\right\} + \left(1 - F\left(\pi_j^*\right)\right) \cdot P_j\right] \end{aligned} \tag{11}$$

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

$$\left.\begin{aligned} &\max_{\{k_{1i}k_{1i}\geq 0, x_{ti}k_{1i}\geq 0\}_{\substack{t=1,2;\\ i=g,b}}} \textstyle\sum_{i=g,b} \upsilon_i \cdot \left(\begin{array}{c} \left(R\left(k_{1i}\right) - x_{1i}\right) \\ +\delta \cdot \left[F\left(\pi_i^*\right) \cdot \left(R\left(k_i^*\right) - x_{2i}\right) + \int_{\pi_i^*}^{\bar{I}} \left(I - P_i\right) \cdot dF\left(I\right)\right] \end{array}\right) \\ &s.t.\ (10),\ (11),\ \text{and } P_i = \max\left\{0, x_{2i}^D - \theta_i \cdot k_i^*\right\}, \end{aligned}\right\} \tag{D}$$

where the last equality follows from equation (7) and Proposition 3.

5.3 No-distortions separating contract

We solve the contracting problem following the same three steps as in the case of zero penalties. We first assume that there are no countervailing incentives and solve for the optimal contract in this case, which happens to be the no-distortions separating contract C^{ND} defined above. Then, we show the conditions under which there are no countervailing incentives. Finally, we characterize the optimal contract when there are countervailing incentives.

As usual, bad GPs' participation and good GPs' incentive compatibility constraints must bind. Therefore, it follows from Proposition 3 that the default penalty with a bad type is $P_b = 0$, as bad GPs do not get any informational rent. Hence, it follows from equation (7) and Proposition 3 that bad GPs' second-period fees are given by $x_{2b}^D = \theta_b \cdot k_b^*$. Using the bad GP's intertemporal participation constraint (equation (10)), we can also obtain the bad GPs' first-period fees, which are $x_{1b}^D = \theta_b \cdot k_{1b}^D$.

As a result, a good GP impersonating a bad one will continue running the fund in the second period if the LP does not opt out the partnership, since $\max\left\{0, x_{2b}^D - \theta_g \cdot k_b^*\right\} = \Delta\theta \cdot k_b^*$. Given that partnership exists are efficient, as shown in Corollary 3, it follows that a good GP must be granted an informational rent of $\Pi_g\left(C_g^D\right) = \Delta\theta \cdot \left(k_{1b}^D + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right)$, as the LP will continue onto the second period with probability $F\left(\pi_b^*\right)$.

Suppose now that the good GP's informational rent were to be paid in full in the second period (we will come back to this point later), i.e., $\Pi_{2g}^D = \Pi_g^D$. Then, using equation (6) and Proposition 3, we have that $\delta \cdot P_g = \Pi_g^D$. Also, the fees perceived by good GPs are given by $x_{1g}^D = \theta_g \cdot k_{1g}$ (i.e.,no informational rent in the first period) and $x_{2g}^D = \theta_g \cdot k_g^* + P_g$ (since the default penalty is equal to the second-period informational rent). We can therefore rewrite the LP's problem as:

$$\max_{\{k_{1i}\geq 0\}_{i=g,b}} \sum_{i=g,b} v_i \cdot \left(\left(R\left(k_{1i}\right) - \theta_i \cdot k_{1i}\right) - \Pi_i^D + \delta \cdot \left(F\left(\pi_i^*\right) \cdot \left(R\left(k_i^*\right) - \theta_i \cdot k_i^*\right) + \int_{\pi_i^*}^{\overline{I}} I \cdot dF\left(I\right) \right) \right),$$

where $\Pi_g^D = \Pi_g\left(C_g^D\right)$ and $\Pi_b^D = 0$, since the bad GP does not get any informational rent.

The FOCs (w.r.t. k_{1i}) of this (concave) problem yield the first period investment levels, which are are given by $k_{1g}^D = k_g^*$ and $k_{1b}^D = k_b^S$, respectively. Hence, as long as the bad GP's incentive compatibility constraint does not bind, the no-distortions separating contract C^{ND}

is implementable. Observe that, again, a good GP would receive an informational rent of $\Pi_g^D = \Pi_g\left(C^{ND}\right) \equiv \Delta\theta \cdot \left(k_b^S + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right)$.

Let us now explore the conditions under which there are no countervailing incentives. Suppose that the no-distortions separating contract is offered. Then, under the assumption that the entire informational rent is paid in the second period, we can write the bad GP's incentive compatibility constraint as:

$$\Pi_b\left(C_g^{ND}\right) = -\Delta\theta \cdot k_g^* + \delta \cdot \left[F\left(\pi_g^*\right) \cdot \max\left\{P_g - \Delta\theta \cdot k_g^*, 0\right\} + \left(1 - F\left(\pi_g^*\right)\right) \cdot P_g\right]. \tag{12}$$

The following proposition characterizes the conditions under which equation $\Pi_b\left(C_g^{ND}\right) \leq 0$, so that bad GPs do not have incentives to impersonate good GPs and the no-distortion contract is therefore incentive-compatible.

Proposition 4 (Range of No-Distortions with Default Penalties) *The no-distortions separating menu of contracts C^{ND} is incentive-compatible with optimally-set penalties if:*

(i) $F\left(\pi_b^*\right) \leq F_b^{CI}$, as defined in Corollary 1.

(ii) Or $F\left(\pi_b^*\right) > F_b^{CI}$ and $F\left(\pi_b^*\right) \geq F_b^D \equiv \frac{\delta \cdot k_g^* - k_b^S}{\delta \cdot k_b^*}$.

(iii) Or $F\left(\pi_b^*\right) > F_b^{CI}$ and $F\left(\pi_g^*\right) \geq F_g^D \equiv 1 - \frac{k_g^*}{k_b^S + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*}$.

Proof. See Appendix. ■

The intuition behind the previous proposition, which is formalized in the proof, goes as follows. Item (i), i.e., $F\left(\pi_b^*\right) \leq F_b^{CI}$, corresponds to the case analyzed above, in which there are no countervailing incentives in the case of zero penalties because the good GP's informational rent is not large enough to attract bad GPs. Hence, the no-distortion menu of contracts is incentive-compatible. We have seen in Proposition 2 that investments must be distorted away from second-best in contracts with zero penalties when $F\left(\pi_b^*\right) > F_b^{CI}$. However, when optimal penalties are included in the contract, the no-distortions menu of contracts is incentive-compatible in a wider range.

Let $F(\pi_b^*) > F_b^{CI}$. Item (ii) corresponds to the case $F_b^D \leq F(\pi_b^*)$, in which the likelihood that the partnership continues onto the second period with a bad type is large. Consequently, the informational rent that must be paid to the good GP to induce separation, which includes the expected value of mimicking the bad type in the second period, is also large. In particular, since the entire informational rent is paid in the second period, running the fund would entail a net profit of $\Pi_{2g}^D - \Delta\theta \cdot k_b^* \geq 0$. Hence, in this case, a bad GP would be willing to run the fund for two periods if called upon by the LP. Therefore, the bad GP would effectively face the same incentives as in a pure repetition of a static contract. Thus, since he is not willing to mimic the good type in a static setting, he will not be willing to do it in a two-period setting either.

Item (iii) corresponds to a case in which the probability of continuation with a good GP is large. In this case, the prospects for a bad GP to cash the good GP's informational rent through the default penalty are so small so as not to be willing to impersonate a good GP.

Figure 5.3 depicts the thresholds F_b^{CI}, F_b^D and F_b^D for a particular set of parameters.[21] Recall from above that $F_b^{CI}\left(\delta^{CI}\right) = F(\pi_b^*)$ and define $F_g^D\left(\delta^D\right) = F\left(\pi_g^*\right)$. For $\delta < \delta^{CI}$, we have that $F_b^{CI}(\delta) < F(\pi_b^*)$, so that the no-distortions menu of contracts is incentive-compatible in that range. For $\delta^{CI} < \delta < \delta^D$, where we have that $F_b^{CI}(\delta) > F(\pi_b^*)$ and also that $F_g^D(\delta) < F\left(\pi_g^*\right)$. In this region, contracts with zero penalties introduce investment distortions, which can be avoided by optimally setting default penalties.

[21] In general, the solid blue and the dashed black line intersect for $\delta = 1$, since $F_b^D|_{\delta=1} = F_b^{CI}|_{\delta=1} = \frac{k_g^* - k_b^S}{k_b^*}$. Figure 5.3 illustrates a case in which we also have $\frac{k_g^* - k_b^S}{k_b^*} < 1$. In Laffont and Tirole (1988), there are countervailing incentives whenever $k_g^* < k_b^* + \delta \cdot k_b^S$. Hence, this case corresponds to a framework in which there are countervailing incentives in Laffont and Tirole (1988) for some $\delta < 1$.

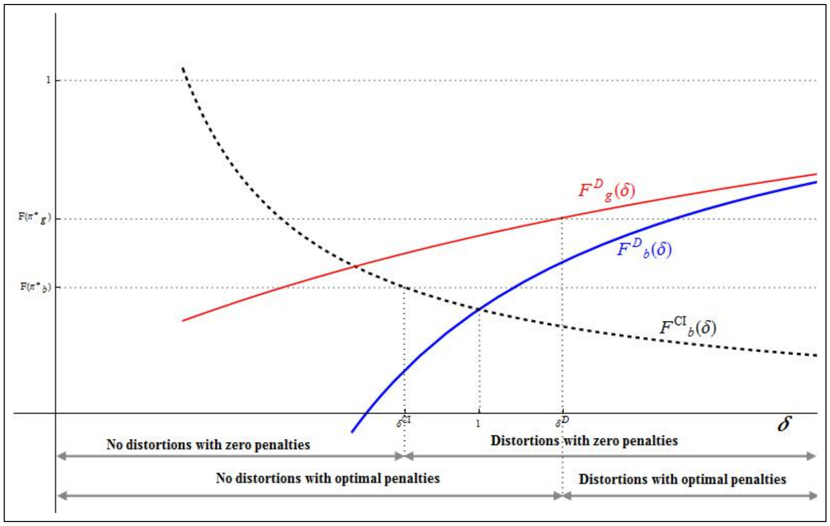

Figure 5.3: Distortions in contracts with zero and with optimal default penalties.

Proposition 4 establishes a role for penalties in two-period contracts. In order to preclude a bad GP from pretending he is a good type, the LP may defer the payment of the informational rent until the second period, so that the bad GP would have to run a large fund for two periods in order to cash this informational rent. However, if the LP could interrupt the partnership without incurring any default penalty, this deferred payment would not be credible, for the LP could potentially call for a partnership interruption and renegotiate the terms of the agreement. In this case, the LP would have to pay this informational rent in the first period and distort investment levels away from the second-best static contracts in order to make the good GP's contract unattractive to a bad GP, as seen in Proposition 2. The role of penalties, by credibly deferring payments to the second period, is that they insulate good GPs' informational rents from partnership terminations without having to make the fee payment up front, as in the case without penalties. In a certain range, deferring the payment of management fees does, therefore, prevent bad GPs from mimicking good GPs without the need of further distortions. Below, we analyze the size of distortions when the no-distortion contract is not incentive-compatible in contracts with optimal fees. Before that, we address the role of deferring fees.

5.4 Optimal deferral of fees

We have assumed above that the entire informational rent was deferred to the second period. Here, we provide a justification of the reason why postponing part–or all– of the informational rent to the second period is optimal. A glance at the ICC constraint in (11) reveals that paying the entire informational rent in the second period relaxes the constraint. The left-hand side of the inequality stands for type$-i$'s intertemporal informational rent, so that any shift of rents across periods, keeping the intertemporal rent 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 share of the informational rent paid in the second period. Hence, deferring informational rents to the second period must be optimal, at least in a weak sense, as it helps incentives. Notice that deferring payments to GPs does not affect the LP's intertemporal earnings.

In the range in which contracts with zero penalties do not introduce distortions there is no need to defer payments to the second period. However, in the range in which penalties eliminate distortions which would otherwise be present, at least a part of the intertemporal informational rent will have to be paid in the second period. Indeed, the very reason why penalties eliminate distortions in certain ranges is precisely because they constitute a mechanism to reduce the incentives to mimic good LPs in the first period by requiring that the fund be run in the second period in order to cash the rent. For the cases in which contracts with optimal penalties prescribe some investment distortion, which we will analyze below, optimality requires that the entire informational rent be deferred to the second period. Even so, some investment distortion is needed to prevent bad GPs from impersonating good GPs. We summarize this discussion in the following proposition.

Proposition 5 (Intertemporal Allocation of Informational Rents) *The optimal intertem-*

poral allocation of informational rents is as follows:

(i) If $F(\pi_b^*) \leq F_b^{CI}$, any intertemporal allocation of informational rents leads to optimal investments.

(ii) If both $F(\pi_b^*) > F_b^{CI}$ and either $F(\pi_b^*) \geq F_b^D$ or $F\left(\pi_g^*\right) \geq F_g^D$, the optimal payment scheme prescribes that at least some part of intertemporal rents be paid in the second period.

(iii) If $F_b^{CI} < F(\pi_b^*) < F_b^D$ and *$F\left(\pi_g^*\right) < F_g^D$, the optimal payment scheme prescribes that the entire intertemporal rents be paid in the second period.*

5.5 Investment distortions

We now analyze the remaining case, i.e., when $F_b^{CI} < F(\pi_b^*) < F_b^D$ *and* $F\left(\pi_g^*\right) < F_g^D$.[22] We will see that, although some investment distortion is required, distortions are strictly smaller than in contracts with zero penalties.

In this case, it follows from the bad GP's incentive compatibility constraint (equation (12)) that $\Pi_b\left(C_g^{ND}\right) > 0$. Hence, the bad GP's incentive compatibility constraint is not satisfied if $k_{1b}^D = k_b^S$ and $k_{1g}^D = k_g^*$. By construction of F_b^D, a bad GP impersonating a good one in the first period will not run the fund in the second period. Hence, we can write his gain from impersonating a good GP as follows:

$$\Pi_b(C_g) = -\Delta\theta \cdot k_{1g} + \delta \cdot \left(1 - F\left(\pi_g^*\right)\right) \cdot P_g.$$

Recognizing that $\delta \cdot P_g = \Delta\theta \cdot (k_{1b} + \delta \cdot F(\pi_b^*) \cdot k_b^*)$, we can then write the LP's problem at time $t = 0$ as:

$$\left.\begin{array}{lll} \max_{k_{1b}\geq 0} & \sum_{i=g,b} \upsilon_i \cdot & \left(\begin{array}{c} (R(k_{1i}) - \theta_i \cdot k_{1i}) - \Pi_i^D(k_{1b}) \\ +\delta \cdot \left(F(\pi_i^*) \cdot (R(k_i^*) - \theta_i \cdot k_i^*) + \int_{\pi_i^*}^{\overline{I}} I \cdot dF(I)\right) \end{array}\right) \\ s.t & k_{1g} = \left(1 - F\left(\pi_g^*\right)\right) \cdot (k_{1b} + \delta \cdot F(\pi_b^*) \cdot k_b^*) & \end{array}\right\}.$$

[22] Observe that $F_b^{CI} < F_b^D$ if and only if $\delta > 1$.

The following proposition characterizes first-period investment levels, while the subsequent one highlights its implications in terms of contract distortions.

Proposition 6 (Optimal Separating Contracts with Default Penalties) *Assume that $F_b^{CI} \leq F(\pi_b^*) \leq F_b^D$ and $F(\pi_g^*) < F_g^D$. Then,*

(i) First-period investment levels satisfy the following conditions:

$$k_{1g}^D = \left(k_{1b}^D + \delta \cdot F(\pi_b^*) \cdot k_b^*\right) \cdot \left(1 - F\left(\pi_g^*\right)\right)$$

and

$$\upsilon_b \cdot R'\left(k_{1b}^D\right) + \upsilon_g \cdot R'\left(k_{1g}^D\right) \cdot \left(1 - F\left(\pi_g^*\right)\right) = \theta_b - \upsilon_g \cdot \theta_g \cdot F\left(\pi_g^*\right).$$

(ii) Second-period investments are efficient, i.e., $k_{2i}^D = k_i^*$.

(iii) Fees are given by $x_{1i}^D = \underbrace{\theta_i \cdot k_{1i}^D}_{Management\ Fees}$ in the first period, and

$$\begin{aligned} x_{2g}^D &= \underbrace{\theta_g \cdot k_g^*}_{Management\ Fees} + \underbrace{\frac{\Delta\theta}{\delta} \cdot \left(k_{1b}^D + \delta \cdot F(\pi_b^*) \cdot k_b^*\right)}_{Transaction\ Fees} \\ x_{2b}^D &= \underbrace{\theta_b \cdot k_b^*}_{Management\ Fees}, \end{aligned}$$

in the second period.

(iv) The LP extracts all the surplus from a bad GP and pays an informational rent $\Pi_g^D \equiv \Delta\theta \cdot \left(k_{1b}^D + \delta \cdot F(\pi_b^*) \cdot k_b^*\right)$ to a good GP.

In order to highlight the role of penalties in private equity contracts, we compare the investment distortions of contracts with optimally-set penalties to those with zero penalties. As the contract distortion depends on the probability that a partnership with a good type will continue after the first period, we carry out comparative statics in which this probability varies. To do so, we fix all the parameters of the model –including $F(\pi_b^*)$, which is set in the range $F_b^{CI} < F(\pi_b^*) < F_b^D$– and we let the distribution function $F(\cdot)$ vary so that

$F\left(\pi_g^*\right)$ varies from $F\left(\pi_b^*\right)$ –its lower bound– to 1.[23] For ease of comparison, section (ii) of the proposition includes the result from Proposition 4.

Proposition 7 (Investment Distortions with Default Penalties) *Assume that $F_b^{CI} < F\left(\pi_b^*\right) < F_b^D$. Then, contracts with optimally-set penalties are such that:*

(i) If $F\left(\pi_g^*\right) < F_g^D$, first-period investments are distorted away from second-best static investment levels, albeit less than in contracts with zero penalties, that is:

$$k_{1b}^Z < k_{1b}^D < k_b^S < k_g^* < k_{1g}^Z < k_{1g}^D.$$

Moreover, the gaps $k_b^S - k_{1b}^D$ and $k_{1g}^D - k_g^*$ shrink as $F\left(\pi_g^*\right)$ increases, getting arbitrarily close to zero as $F\left(\pi_g^*\right)$ approaches F_g^D.

(ii) If $F\left(\pi_g^*\right) \geq F_g^D$, first-period investments correspond to the second-best static investment levels, while investments in contracts with zero penalties are distorted away from the second-best static investment levels. That is:

$$k_{1b}^Z < k_{1b}^D = k_b^S < k_g^* = k_{1g}^D < k_{1g}^Z.$$

Proof. See Appendix. ■

Default penalties constitute a gamble for a bad GP willing to impersonate a good GP. As the LP's outside option becomes more attractive, the likelihood of cashing the default penalty (and therefore of cashing the informational rent without having to run the fund) increases. Hence, the incentives to mimic a good GP are enhanced the lower the probability that the partnership continues to the second period with a good GP. The profit-maximizing contract distorts the investment level so as to make the good GP's contract unattractive to the bad GP. However, this distortion is always smaller than in the case with no penalties. In the case with no penalties, since informational rents are paid up front, informational rents would be cashed by a bad GP for certain. Hence, the good GP compensation package is more attractive, and

[23] Observe that, whenever $F_b^{CI} < F\left(\pi_b^*\right)$, it follows that $F_g^D < 1$.

so distortions have to be larger.

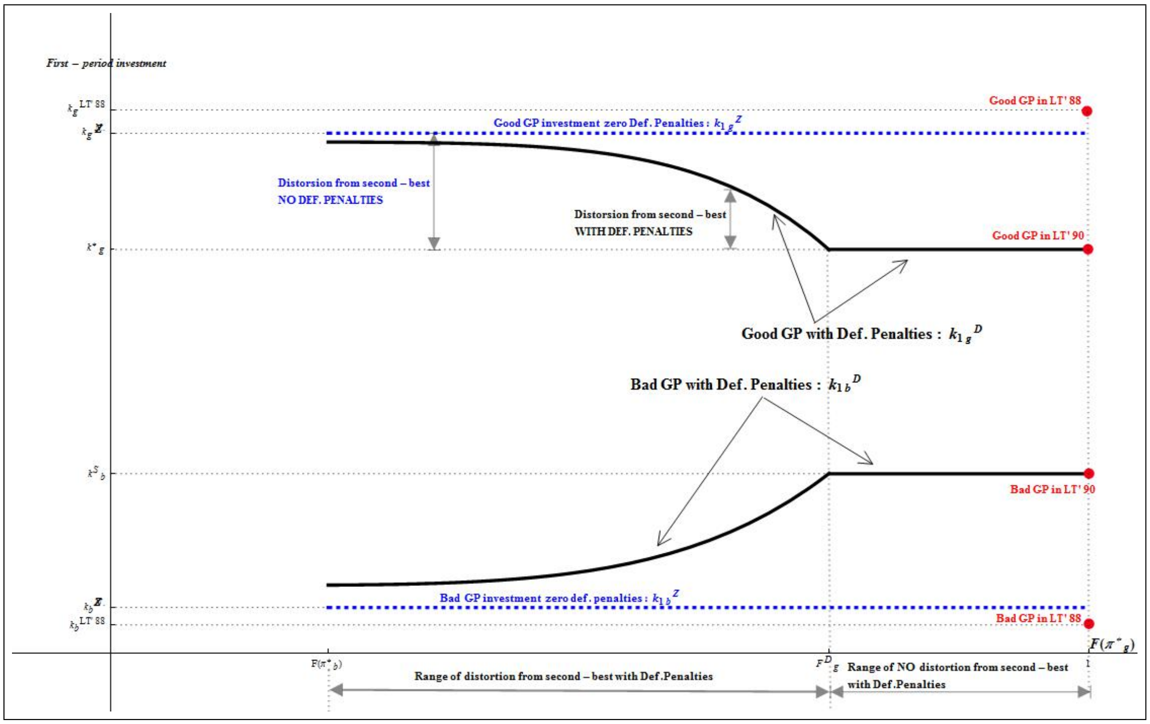

Figure 5.5: Distortions in contracts with default penalties.

Figure 5.5 illustrates this proposition. Given some $F\left(\pi_b^*\right) \in \left(F_b^{CI}, F_b^D\right)$, we represent the probability of continuation with a good agent $F\left(\pi_g^*\right)$ in the horizontal axis and first-period investment levels in the vertical axis. The blue dotted horizontal lines represent first-period investments for a good GP (upper line) and a bad GP (lower line) for contracts with zero default penalties. From Proposition 2, these investment levels are distorted away from the second-best static investment levels. For the sake of comparison, we include in the figure the contracts in Laffont and Tirole (1988), which correspond to a probability 1 of continuation. In our setting, the LP may opt out of the partnership with some probability. Hence, the up-front informational rent that must be paid to a good GP is smaller. Therefore, the distortion away from second-best static levels is smaller in our framework. The black solid lines represent first-period investments for a good GP (upper line) and a bad GP (lower line) for contracts with optimally-set default penalties. As shown in Proposition 7, for $F\left(\pi_g^*\right) < F_g^D$ investments must be distorted, albeit less so than in contracts with zero penalties. Hence, the black solid lines are closer to second-best investment levels than the blue dotted lines. For $F\left(\pi_g^*\right) > F_g^D$, the no-distortion contract is incentive-compatible with optimally-set penalties. Hence, there is no distortion in this range. Again, for comparison purposes, we depict the corresponding contracts in Laffont and Tirole (1990), which correspond to a probability 1 of continuation and no distortions from second-best static investment levels.

5.6 Full-commitment versus contracts with no-commitment and default penalties

We now turn into the comparison between full-commitment contracts, which we analyzed in section 3.3, and the contracts with no-commitment and default penalties analyzed above.

Recall that the full-commitment two-period menu of contracts consists of a repetition of second-best static contracts. Hence, the LP would obtain a profit of $\Pi_L\left(C^{FC}\right) = (1+\delta)\cdot\Pi_L^S$ if that contract could be offered. Now, consider the case of contracts with default penalties. For notational simplicity, let $q \equiv F\left(\pi_g^*\right)$ the probability of continuation with a good GP in the second period. In the first period, contracts with default penalties yield either second-best static investment levels or some distorted investments. We can write the first-period LP's profit function as $\Pi_L^S - D(q)$, where $D(q)$ stands for the reduction in first-period profits due to the investment distortion. In light of the previous proposition, we have that $D(\ q) = 0$ for either $q \geq F_g^D$, or $F(\pi_b^*) \leq F_g^{CI}$ or $F(\pi_b^*) \geq F_g^D$. In the remaining region, $D(q)$ is strictly decreasing in q. In the second period, depending on the realization of the outside value I, the partnership leads to either first-best separating contracts or continuation at first-best with good GPs or partnership interruption. We can write LP's profits from a contract with penalties as:

$$\Pi_L\left(C^D\right)=\begin{cases}\Pi_L^S-D\left(q\right)+\delta\cdot\left(\upsilon_g\cdot\left(\pi_g^*-P_g\right)+\upsilon_b\cdot\pi_b^*\right) & \text{if } I<\pi_b^* \\ \Pi_L^S-D\left(q\right)+\delta\cdot\left(\upsilon_g\cdot\left(\pi_g^*-P_g\right)+\upsilon_b\cdot I\right) & \text{if } \pi_b^*\leq I\leq\pi_g^* \\ \Pi_L^S-D\left(q\right)+\delta\cdot\left(I-P_g\right) & \text{if } I>\pi_g^*\end{cases},$$

where P_g stands for the informational rent that the LP must pay to good GPs regardless of whether the partnership extends to the second period or not. There are two advantages of full-commitment contracts over contracts with default penalties. On the one hand, with full-commitment contracts there is no first-period investment distortions over second-best static investment levels. This advantage applies as long as we are in the distortions region. Otherwise, both contracts lead to the same (second-best static) investment levels.

On the other hand, the ratchet effect leads to potentially large informational rents in the case of contracts with default penalties, because good GPs must be compensated. Notice, however, that with random interruption of partnerships, the informational rents associated with second-period mimicking may even be smaller in the case of default penalties. In particular, they will be smaller whenever $F\left(\pi_b^*\right)\leq\frac{k_b^S}{k_b^*}$. Hence, we have that whenever $F\left(\pi_b^*\right)\leq\min\left\{F_b^{CI},\frac{k_b^S}{k_b^*}\right\}$, contracts with default penalties yield strictly higher profits to LPs than full-commitment contracts.

In addition, contracts with default penalties confer the LP flexibility to enjoy a potentially high outside value. Hence, even though contracts with default penalties may induce some first-period investment distortions and pay larger informational rents, the prospects of good outside opportunities make contracts with default penalties particularly desirable. In Figure 5.6 we represent LP's second-period profits as a function of potential realizations of the outside value I. The solid black line represents LP's profits under a contract with default penalties, while the dotted red line stands for those corresponding to a full-commitment contract. The figure depicts a case in which the informational rent paid in contracts with default penalties is larger. Therefore, the solid black line lies below the dotted red line for all $I<\pi_b^*$, in which we would have continuation with both types of GPs under both contracts. The value of flexibility to enjoy outside options conferred by contracts with penalties is captured by the segments in which the partnership is interrupted with some of the GP types. In particular, for sufficiently high realizations of the outside option, LP's profits are higher under contracts with default penalties. Hence, contracts with default penalties would yield higher profits for outside value distributions with a large upside.

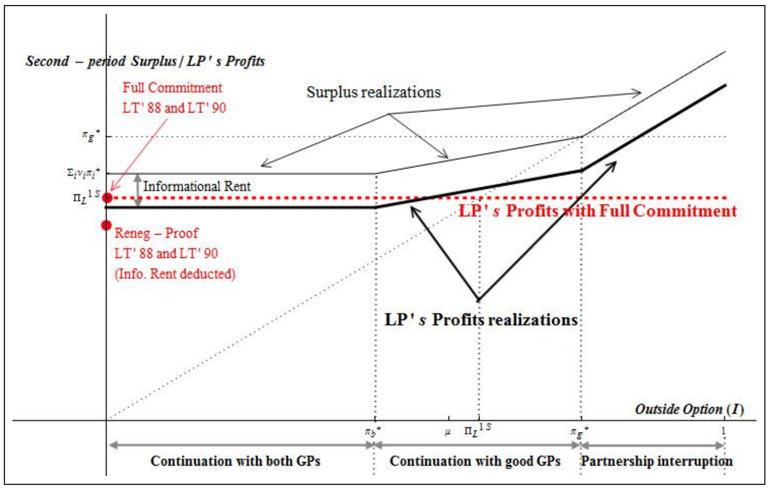

Figure 5.6: Second-period LP's Profits with (FC) and (D) contracts

6 Pooling versus separating contracts

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 $C^P = \left\{C_g^P, C_b^P\right\}$ and address the effects of the LP's second-period outside option in the design of the contract.

6.1 Pooling contracts

Consider the case in which the profit-maximizing two-period contract entails first-period pooling. Since first-period investment levels do not interfere with those of the second period, a pooling contract must be optimal within the class of static pooling contracts. Hence, the size of the pooling fund in the first period would be $k_{1i} = k_b^*$.[24]

With a pooling contract, the LP's beliefs about the type of the GP are not updated after first-period outcomes are realized. Then, 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 would continue onto the second period with any GP must therefore be the profit-maximizing static contract. Hence, if the partnership extends onto the second period with both types of GPs, we would necessarily have types' separation in the second period, with funds of size $k_{2g}^P = k_g^*$ and $k_{2b}^P = k_b^S$, as in the second-best static contract.

Nonetheless, the profit-maximizing second-period contract may entail the exclusion of the bad type for some realizations of the LP's outside option. While the conditions that we have imposed on the marginal returns for low investment levels ensure that shutting down a bad GP is not optimal in the first period, the possibility of enjoying her outside option opens up a window for optimally shutting down a bad GP in the second period. The LP may propose a unique contract with $k_{2g}^P = k_g^*$, which only the good GP would accept, and enjoy his outside option in case the contract is not accepted. This contract would lead to an expected payoff of $\upsilon_g \cdot \pi_g^* + \upsilon_b \cdot I$, which would exceed Π_L^S whenever $I > I^P$, where:

$$I^P \equiv \pi_b^S - \frac{\upsilon_g}{\upsilon_b} \cdot \Delta\theta \cdot k_b^S.$$

Observe first that I^P is positive, which follows from the assumption that $\lim_{k \to 0} R'(k) \cdot k = 0$. Hence, while static profit-maximizing with a zero outside option for the LP prescribes contracting with a bad GP for certain, it is optimal for the LP to only continue the partnership with a good GP when his outside option is large enough. Moreover, since $\upsilon_g \cdot \pi_g^* + \upsilon_b \cdot I^P = \Pi_L^S$, it follows that $I^P < \pi_b^*$. Hence, partnerships with a bad GP would be broken even when it would be efficient to continue, namely for realizations $I \in \left(I^P, \pi_b^*\right)$.

This inefficiency stands in sharp contrast to profit-maximizing contracts with first-period

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

separation, in which partnerships extend to the second period if and only if it is efficient to do so. When there is separation in the first period, the LP can extract all the second-period surplus. Hence, she does only opt out of the partnership when the surplus that can be generated is lower than her outside option. If, on the contrary, the contract entails pooling in the first period, unveiling the GP's type requires paying an informational rent to a good GP, so he cannot appropriate the entire second-period surplus. This informational rent reduces the gains from contracting with a bad GP, which makes the outside option relatively more valuable.

Finally, the LP can interrupt the partnership regardless of the type of the GP, in which case he would enjoy his outside option I. He will opt out the relationship whenever $I \geq \pi_g^*$.

Trivially, optimal penalties must be zero in any contract with types pooling in the first period. If there were a positive penalty to interrupt the partnership, it would have to be paid to both types of the GP. But that would entail a positive transfer to the bad type without helping incentives, so that any such penalty would strictly reduce the LP's profits.

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

Proposition 8 (Optimal Pooling Contracts) *The optimal pooling contract is as follows:*

(i) First-period investment and fees are given by $k_1^P = k_b^*$ *and* $x_1^P = \theta_b \cdot k_b^*$*, respectively.*

(ii) Second-period investment levels are second-best static, i.e., $k_{2g}^P = k_g^*$ *and* $k_{2b}^P = k_b^S$ *(conditional on continuation with both types).*

(iii) Second-period fees are effectively imposed by the LP once her outside option has been realized, being $x_{2g}^P(I) = \theta_g \cdot k_g^*$ *if* $I \in \left[I^P, \pi_g^*\right]$ *and* $x_{2g}^P(I) = \theta_g \cdot k_g^* + \Delta\theta \cdot k_b^S$ *if* $I < I^P$ *for the good GP and* $x_{2b}^P = \theta_b \cdot k_b^S$ *for the bad GP.*

(iii) Default penalties are zero.

Observe that, since default penalties are zero (so that the LP can renege on any previous agreement at no cost), whenever the outside value is such $I \in \left[I^P, \pi_g^*\right]$, the LP will offer a single contract, with investment $k_{2g}^P = k_g^*$ and fees $x_{2g}^P(I) = \theta_g \cdot k_g^*$. This contract will only be taken by good GPs, who will break even. If $I < I^P$, second-best contracts will be offered. Hence, in this case, the rent of good GPs will increase, as optimal contracting with both types entails separation, which in turn requires an informational rent to the good GP.

6.2 Profit-maximizing two-period contracts

Although under Assumption 1 separating contracts are always preferred to pooling contracts in static settings, there is a standard trade-off when there is repetition of contracts, as identified in the classical treatments of two-period adverse selection models. Notwithstanding, the LP's outside option plays an important role in the determination of this threshold.

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 he has to give up larger informational rents to induce separation in the first period than in the full-commitment setting, in which static second-best investment levels would be implemented in both periods. These extra informational rents are larger the longer the duration of the second period. Therefore, when the second period extends for a sufficient length, the LP will not be willing to incur the cost of inducing separation, Instead, he would prefer to give up some efficiency in the first period in exchange for a reduction of informational 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.

In order to grasp the intuition behind the effect of the LP's outside option in the determination of this threshold, consider a situation in which the outside option prospects are very high, so that any partnership would be likely to be broken before the second period. This context would be very similar to a static contracting framework. Consequently, we would expect that first-period separation would dominate pooling contracts, regardless of the second-period duration. The opposite situation would arise if, for instance, it was very likely that the outside option fell into the range in which the partnership would continue with either type –for instance, if the outside option was (close to) zero. Then, a pooling contract may always be preferred to separation when the second-period duration is long enough.

In order to formalize the implications of the LP's outside option on the comparison between separating and pooling contracts, consider a family of probability distribution functions $\{F_\alpha(\cdot) : \alpha \in [0,1]\}$ on the LP's outside option with the following properties:

(i) First-order stochastic domination in the range $[0, \pi_b^*)$: For any $\alpha' < \alpha''$, $F_{\alpha'}(I) < F_{\alpha''}(I)$ for all $I < \pi_b^*$.

(ii) Border conditions: For all $I < \pi_b^*$, $F_0(I) = 0$ and $F_1(I) = F(\pi_b^*)$.

(iii) For any α, $F_\alpha(\pi_b^*) = F(\pi_b^*) \leq 1$.

Observe that smaller $\alpha -$ *indexes* signify a lower concentration of probability mass in small values of the LP's outside option, the extreme ($\alpha = 0$) being a distribution with zero mass for all $I < \pi_b^*$ and an mass point at $I = \pi_b^*$. At the other extreme ($\alpha = 1$), we would have a distribution with zero mass for all $I \in (0, \pi_b^*]$ and a mass point at $I = 0$.[25] Hence a function with a smaller α would first-order stochastically dominate a function with a higher α in the range below π_b^*. We leave absolute freedom in the structure of the family for $I > \pi_b^*$, as the trade-off first-period separating versus pooling is unaffected by outside value realizations above π_b^*.[26] The following proposition characterizes the second-period duration threshold that determines whether separation is preferred to pooling.

Proposition 9 (Optimality of Pooling versus Separating Contracts) *Assume that $F(I) > 0$ for some $I < \pi_b^*$. Then:*

(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 $\delta > \delta^P$.

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

[26] Observe that if we construct the family leaving the upper part of the distribution unchanged, that is, $F_\alpha(I) = F(I)$ for all $I > \pi_b^*$, then we would have that a function with a smaller α would first-order stochastically dominate a function with a higher α in the entire range. Consequently, a lower α would be associated with a higher outside value expectation μ_α. We could also keep expectations unchanged across the family, i.e., $\mu_\alpha = \mu$ for all α, which would require that for any $\alpha' < \alpha''$, $F_{\alpha'}(I) > F_{\alpha''}(I)$ for some range(s) within $I > \pi_b^*$.

(ii) Let δ_{α}^{P} be the pooling threshold associated with the distribution $F_{\alpha}(\cdot)$ from the family described above. Then, for any $\alpha' < \alpha''$, we have that $\delta_{\alpha'}^{P} < \delta_{\alpha''}^{P}$, 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. While the result (i) in the Proposition is standard, it is instructive to analyze the advantages of a separating contract over a pooling one in the presence of an outside option in the second period. A separating contract rewards the LP with a higher first-period profit ($\Pi_L^S > \pi_b^*$). However, it entails too large a transfer to also induce separation in the second period.

Observe that second-period pooling and separating contracts only differ in the range $I < \pi_b^*$. For larger realizations of the outside option, second-period contracts prescribe the exclusion of bad GPs (if $I \in \left(\pi_b^*, \pi_g^*\right)$) or of both types of GP (if $I \geq \pi_g^*$), regardless of first-period outcomes. Hence, the comparison of separating versus pooling contracts must only consider events in the range $I < \pi_b^*$. Hence, for distributions with all mass at or above π_b^*, separating contracts would always be preferred to pooling contracts and, hence, we would have that $\delta^P = \infty$.

Let us now consider values below π_b^*. On the one hand, if $I < I^P$, first-period separation leads to second-period efficient contracts (with large transfers), leading to LP's profits given by $\Pi_L^* = \Pi_L^S - \upsilon_g \cdot \Delta\theta \cdot \left(k_b^* - k_b^S\right)$. In this range, first-period pooling induces second-period second-best static contracts. On the other hand, if $I \in \left[I^P, \pi_b^*\right]$, first-period separation also leads to second-period efficient contracts, while first-period pooling induces second-period exclusion of a bad GP, leading to LP profits of $\upsilon_g \cdot \pi_g^* + \upsilon_b \cdot I$.

Observe that, by construction of I^P, it follows that $\Pi_L^* < \Pi_L^S < \upsilon_g \cdot \pi_g^* + \upsilon_b \cdot I$ 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 will dominate separating contracts whenever the second-period duration is sufficiently large, i.e., $\delta^P < \infty$. Observe also that the advantage of pooling versus separating is larger the larger the outside option. Hence, for distributions that concentrate higher amounts of mass in larger values of the outside option, pooling will be relatively more profitable. Hence, if a distribution first-order stochastically dominates another, then the former will have a smaller associated pooling threshold.

7 Empirical predictions

We now lay out the empirical predictions generated by the model. We first draw implications for the case of separation. At the end, we propose a test for distinguishing between the case of separation and that of pooling.

7.1 Default Penalties

The first set of predictions is about the extensive margin in the use of default penalties. Default penalties arise as a response to countervailing incentives. Therefore, they should be present when the conditions for countervailing incentives emerge. The relevant condition is the one required for the implementability of no-distortions renegotiation proof contracts, equation (4), which states that no-distortion contracts are implementable when $k_g^* \geq k_b^S + \delta \cdot F(\pi_b^*) \cdot k_b^*$. When this inequality does *not* hold, countervailing incentives and default penalties arise. For our first set of predictions, we look at what happens when the right-hand side of the inequality increases, thus making the need for default penalties more likely.

Predictions 1-2 *(Extensive Margin) The probability of observing default penalties increases with 1) the probability of continuation with a bad GP ($F(\pi_b^*)$), and 2) the residual life of the fund at the time of exit (or, equivalently, the length δ of the second period).*

Prediction 1 is based on the willingness/ability of an LP to continue providing capital to a fund run by a low-quality GP. The prediction is essentially about the opportunity cost of providing investment in a low-quality fund.

Prediction 2 establishes a link between the probability of observing default penalties and the duration of the second period. When the second period has a larger weight in the overall performance of the fund, then default penalties are more likely to arise. This is an intuitive result: the rents that a good GP forgoes when the LP defaults increase in the length of the second period.

Our second set of predictions is about the intensive margin in the use of default penalties. Proposition 3 states that at the optimum default penalties are set equal to the informational rents of each type of GP. As the information rent of a bad GP is zero, the default penalty for the bad GP is also set to zero. The rent of a good GP is equal to $\Pi_g\left(C_g^D\right) = \Delta\theta \cdot \left(k_{1b}^D + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right)$, which also defines the value of the default penalty for this type of GP. We obtain the following predictions:

Prediction 3-6 *(Intensive Margin) Default penalties 3) are non-negative for good GPs and zero for bad GPs, 4) are higher for larger funds, 5) increase in the probability of continuation with a bad GP ($F\left(\pi_b^*\right)$), 6) increase in the residual life of the fund at the time of exit by the LP (δ).*

Predictions 3 states that only good GPs need to be compensated for a loss of rent when the LP exits the fund, while Prediction 4 states that better GPs run larger funds. Putting these two predictions together, we obtain that penalties should be higher for larger funds. The only analytical academic work on default penalties is that of Litvak (2004), which shows that default penalties are higher (in terms of coefficient of severity) in larger funds. 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 predictions 3 and 4: quality of GPs, size of the fund, and severity of the penalty are all positively correlated.[27]

Prediction 5 positively relates the intensive margin of default penalties to the probability

[27] Litvak's (2004) Table 1 suggests that the coefficient of severity is positively related to a monetary loss for an LP.

of continuation with a bad GP. Combining predictions 2 and 5 we find that the probability of continuation increases both the probability of observing default penalties and their size (conditional on observing them).

Finally, Prediction 6 relates the duration of the second period to the size of the default penalty. In our model, the longer the second-period duration, the higher the default penalty. Litvak (2004) provides evidence that supports this prediction. In particular, Litvak shows that default penalties are positively correlated with the option term, which constitutes a measure of the relative importance of future capital calls.[28]

7.2 Fee structure

The model calls for fees that are proportional to the capital under management, as well as for non-proportional fees. The real-world counterpart of the proportional fees is given by management fees, while non-proportional fees would correspond to transaction fees.

Prediction 7 *(Management Fees) GPs are compensated with a fee that is expressed as a percentage of the size of the fund in each period. The percentage is smaller for better GPs, which also implies that larger funds require lower percentage fees.*

Prediction 7 stems from the fact that GPs are compensated with a fee that is equal to $\theta_g k_g$ and $\theta_b k_b$, respectively for the good and bad GP. Given that $\theta_g < \theta_b$, good GPs should receive a smaller proportional fee. Additionally, given that in equilibrium good GPs run larger funds, proportional fees at larger funds should represent a smaller percentage of fund size than at smaller funds.

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

[28] 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.

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.

The model shows that LPs might also need to pay non-proportional fees that are not paid at the inception of the fund. The natural interpretation of these fees is that they represent transaction fees. Transaction fees are charged by GPs to portfolio firms in connection with the completion of the acquisition for—typically unspecified—advisory services. As discussed in Phalippou, Rauch, and Umber (2015), transaction fees are seldom fully rebated against management fees.[29] Therefore, transaction fees represent charges that GPs impose on the fund and that indirectly affect LPs' profits. In the model, the transaction fees are equal to the rent of the good GP in the second period. When the implementability condition (4) is satisfied, the rent of the good type (and also of the bad type) may be zero in the second period because all rents may be paid up front in the first period. When the implementability condition is not satisfied, there are countervailing incentives and, as a consequence, default penalties. In this case, the rent of a good GP is $\Pi_g\left(C_g^D\right)$, i.e. it is the same as the default penalty.

Prediction 8 *(Transaction Fees) Good GPs are compensated with fees that are non-*

[29] 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."

proportional to capital under management, and that are paid at the start of the second period.

Metrick and Yasuda (2010b) report that the GPs of buyout funds charge transaction fees which vary between 1% and 2% percent of the transaction value. Phalippou, Rauch, and Umber (2015) 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, if not fully rebated), and that larger funds levy larger transaction fees in absolute amounts of dollars.

7.3 Separating versus pooling

Finally we suggest a possible way to test whether the equilibrium may be a separating or a pooling one. In the range of non-distortionary separating equilibria, good GPs are instructed to run a fund at their first-best level of capital in both periods. On the contrary, bad GPs display a certain degree of ratcheting, jumping up from second-best to first-best. First-period pooling requires that both GPs invest at the level which is first-best for the bad GP. In the second period, contracts are set at the static second-best for both GPs. Hence, with pooling, in the second-period we observe ratcheting for good GPs, and a reduction in the size of the fund for bad GPs.

In light of these differences, we can make the following prediction.

Prediction 9 *(Separating versus pooling) Capital calls that are equally split are suggestive of a separating equilibrium. A reduction of the size of capital calls over time suggests that the equilibrium is a pooling one. Any other pattern (including increasing capital calls) is inconclusive.*

8 Conclusions

This paper provides a two-period model which describes the interaction between LPs and GPs as a two-period principal-agent adverse selection problem. LPs set the contractual terms and conditions to screen GPs of heterogeneous and unobservable ability. Optimal contracts include default penalties that LPs have to pay if they do not honor a capital call. The model is particularly suited to analyze 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 distort investment size to better screen GPs. The extent of distortion is smaller if default penalties are included in the agreement. The main problem that arises in a setting of two-period adverse selection is that both good and bad GPs (agents) may be tempted to mimic the other type, a phenomenon known as countervailing incentives. To address countervailing incentives, Laffont and Tirole (1988) show the good type needs to be distorted upwards and the bad type downwards. In our setting, this implies that to deal with countervailing incentives, LPs would need to overinvest and overpay good GPs, and underinvest and underpay bad GPs. We show that these distortions can be reduced by introducing default penalties in the contract. Default penalties act as an insurance mechanism for good GPs, and give the contracting parties the ability to postpone 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 fee structure. The common fee structure employed by private equity funds comprises three types of fees: monitoring/management fees, carried interest, and transaction fees. The latter are charged by GPs to portfolio firms and then partially rebated to the GPs (as a discount on management fees due). Our model shows that management and transaction fees are part of an optimal contract. The 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.

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.

Bernardo, A. E., H. Cai, and J. Luo (2004): "Capital Budgeting in Multidivision Firms: Information, Agency, and Incentives," *Review of Financial Studies*, 17(3), 739–767.

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. (1988): "Commitment Through Renegotiation-Proof Contracts with Third Parties," *Review of Economic Studies*, LV, 377–90.

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 (1987): "Comparative Statics of the Optimal Dynamic Incentive Contract," *European Economic Review*, 31, 901–926.

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.

Laffont, J.-J., and D. Martimort (2002): *The Theory of Incentives*. Princeton University Press, Princeton, NJ.

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*, p. hhq020.

——— (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 (2015): "Private Equity Portfolio Company Fees," *Available at SSRN 2702938*.

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.

A Derivation of benchmark contracts

In this section, we provide a derivation of the benchmark contracts that has been laid out in Section 3. Further details can be found in Laffont and Martimort (2002) (Chapters: 2 for static contracts, 8 for full-commitment contracts).

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:

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

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

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

$$x_i^* = \theta_i \cdot k_i^*.$$

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

$$\pi_i^* \equiv R(k_i^*) - \theta_i \cdot k_i^*.$$

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 $\pi_g^* > \pi_b^*$, that is, a good GP can potentially generate a higher surplus.[30]

[30] By strict concavity of R and optimality of the good type's efficient investment level, we have that $\pi_g^* \equiv R\left(k_g^*\right) - \theta_g \cdot k_g^* > R\left(k_b^*\right) - \theta_g \cdot k_b^*$. Also, since $k_b^* > 0$, we have that $\theta_g < \theta_b$ implies that $R\left(k_b^*\right) - \theta_g \cdot k_b^* > R\left(k_b^*\right) - \theta_b \cdot k_b^* \equiv \pi_b^*$.

A.2 Second-best static contracts

Let us from now on consider the asymmetric information case. In a static framework, the LP solves the following optimization program:

$$\left.\begin{array}{lll} \max_{\{k_i\geq 0, x_i\geq 0\}_{i\in\{g,b\}}} & \sum\limits_{i=g,b} \upsilon_i \cdot [(R(k_i) - x_i)] & \\ s.t & x_i - \theta_i \cdot k_i \geq 0 & [PC.i] \\ & x_i - \theta_i \cdot k_i \geq x_j - \theta_i \cdot k_j & [ICC.i] \end{array}\right\},$$

where $(PC.i)$ and $(ICC.i)$ stand for type$-i$'s participation and incentive compatibility constraints, respectively.

This is a standard static adverse selection problem. Whether this program is solved by a pooling or by a separating contract depends on the ex-ante likelihood that the GP is good. In what follows, we provide the conditions for static separating equilibria to be preferable than pooling contracts.

A.2.1 Second-best static separating contracts

In a separating profit-maximizing contract, efficient investment levels for both types of the LP would not be implementable, for the good type would be willing to mimic the bad one. Notice that the bad GP's participation constraint should bind, as giving up rents to this type would only harden the incentive compatibility constraint for the good type. Hence, in a *separating second-best static (S)* contract, we would have that $x_b^S = \theta_b \cdot k_b^S$. Moreover, the good GP's incentive compatibility constraint should bind as well, as otherwise the good GP would be granted an unnecessarily large transfer. Hence, we have that $x_{1g} = \theta_g \cdot k_g^S + \Delta\theta \cdot k_b^S$, so that the good GP should be compensated with an amount $\theta_g \cdot k_g^S$ for the opportunity cost incurred in running the fund, as well as with an informational rent $\Delta\theta \cdot k_b^S$, so that he would not be willing to impersonate the bad type. Hence, the LP's problem would reduce to:

$$\max_{\{k_i\geq 0\}_{i\in\{g,b\}}} \quad \sum_{i=g,b} \upsilon_i \cdot [(R(k_i) - \theta_i \cdot x_i)] - \upsilon_g \cdot \Delta\theta \cdot k_b \quad ,$$

with transfers 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 good type, that is, $k_g^S = k_g^*$ which, as seen above, satisfies:

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

However, the bad type's investment level satisfies:

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

Clearly, we have that $k_b^S < k_b^*$, which follows directly from the concavity of R and the fact that $\theta_b + \frac{\nu_g}{\nu_b} \cdot \Delta\theta > \theta_b$. Observe that a separating contract could induce first-best investment levels, but that this would be too costly: the LP would have to pay an informational fee of $\Delta\theta \cdot k_b^*$ for a good GP not to be willing to impersonate a bad one. The LP can reduce this fee to $\Delta\theta \cdot k_b^S$ by reducing the size of a fund managed by a bad GP. This trade-off between efficient investment and the informational rent given up to the good GP is resolved by distorting the size of a fund managed by a bad GP downward to k_b^S.[31]

A.2.2 Static pooling contracts

Consider now the option of granting a unique pooling contract, so that $k_g = k_b = k^{SP}$ and $x_g = x_b = x^{SP}$. Then, the incentive compatibility constraints are trivially satisfied. Also, since $\theta_g < \theta_b$, the participation constraint for the bad type implies that the good type's participation constraint is also satisfied. Hence, since giving up rents to the bad type would be suboptimal, it follows that $x^{SP} = \theta_b \cdot k^{SP}$. Therefore, the LP's problem would reduce to:

$$\max_{k\geq 0} \quad R(k) - \theta_b \cdot k.$$

Since 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} would then satisfy:

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

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

A.2.3 Profit-maximizing static separating contracts

Offering a pair of contracts with investment levels $\left\{k_g^*, k_b^S\right\}$ would grant the LP an expected payoff of:

$$\Pi_L^S \equiv \upsilon_g \cdot \left(\pi_g^* - \Delta\theta \cdot k_b^S\right) + \upsilon_b \cdot \pi_b^S,$$

where $\pi_b^S \equiv R\left(k_b^S\right) - \theta_b \cdot k_b^S$ stands for the surplus generated by a bad GP when he invests the amount k_b^S prescribed by a separating static contract.

The LP's payoff from a static pooling contract would be given by:

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

Effectively, in a pooling contract the LP would get the same profits as if the GP was known to be the bad type, for all the extra surplus generated by a good GP would be fully appropriated by the agent through the informational rent $\Delta\theta \cdot k_b^*$.

Whether the profit-maximizing static contract entails pooling or types separation depends crucially on the ex-ante likelihood υ_g that the GP is good. A menu of separating contracts specifies an efficient investment level k_g^* and a (relatively small) transfer $\Delta\theta \cdot k_b^S$ to a good GP. Pooling contracts, on the contrary, prescribe an inefficiently low level of investment k_b^* for a (relatively large) transfer $\Delta\theta \cdot k_b^*$ to a good GP. Hence, the more likely the GP is good, the better a separating contract.[32]

[31] Offering a menu of separating contracts entails an informational transfer to good GPs. The LP could avoid this extra transfer by offering a unique (efficient) contract $\left\{k_g^*, x_g^*\right\}$ that only a good agent would accept. However, our assumption that $\lim_{k\to 0} R'(k) \cdot k = 0$ ensures that this would not be optimal, even if the likelihood that the GP is good was very large. In Appendix B we show why this is the case.

[32] However, a pooling contract may never be profit-maximizing. Although a pooling contract prescribes an efficient level of investment for a bad GP, it also specifies a large transfer to a good one. As the likelihood that the GP is bad increases, separating contracts approach pooling contracts, as the investment distortion for the bad GPr gets arbitrarily close to zero. Depending on parameters, we may have that separating contracts

Assumption 1 guarantees that the optimal static contract is separating.

A.3 Full-commitment two-period contracts

The LP's *Full Commitment (FC)* problem consists of designing a menu of contracts $C^{FC} = \{C_g^{FC}, C_b^{FC}\}$ such that both parties must abide by the terms of the partnership for both periods. A two-period contract specifies investment levels and transfers for each period, potentially as a function of all available history. The LP's problem is given by:

$$\left.\begin{array}{lll} \max_{\{k_{1i}\geq 0, x_{1i}\geq 0\}_{\substack{t=1,2\\ i=g,b}}} & \sum_{t=1,2}\delta_t \cdot \sum_{i=g,b} \upsilon_i \cdot [(R(k_{ti}) - x_{ti})] & \\ s.t & \sum_{t=1,2}\delta_t \cdot (x_{ti} - \theta_i \cdot k_{ti}) \geq 0 & [PC.i] \\ & \sum_{t=1,2}\delta_t \cdot (x_{ti} - \theta_i \cdot k_{ti}) \geq \sum_{t=1,2}\delta_t \cdot (x_{tj} - \theta_i \cdot k_{tj}) & [ICC.i] \end{array}\right\}$$

First, observe that the contract cannot possibly pin down each period transfer, but simply an intertemporal transfer $\overline{x}_i \equiv \sum_{t=1,2}\delta_t \cdot x_{ti}$ for each potential type of the GP. As above, both the good type's incentive compatibility and the bad type's participation constraints are binding. Hence, we can write the previous maximization problem as:

$$\max_{\{k_{1i}\geq 0\}_{t=1,2;i=g,b}} \quad \sum_{t=1,2}\delta_t \cdot \sum_{i=g,b} [\upsilon_i \cdot [R(k_{ti}) - \theta_i \cdot k_{ti}] - \upsilon_g \cdot \Delta\theta \cdot k_{tl}],$$

and intertemporal transfers given by $\overline{x}_g = \sum_{t=1,2}\delta_t \cdot (\theta_g \cdot k_{tg} + \Delta\theta \cdot k_{tb})$ and $\overline{x}_b = \sum_{t=1,2}\delta_t \cdot \theta_b \cdot k_{tb}$, respectively.

The first-order conditions are necessary and sufficient for a solution to this concave program. Then, we have that the optimal contract lasting for two periods consists of a repetition of the second-best static problem, that is:

$$k_{tg}^{FC} = k_g^*, \text{ for } t = 1, 2,$$

and

$$k_{tb}^{FC} = k_b^S, \text{ for } t = 1, 2.$$

B Excluding the bad type

In this section, we provide an analysis of the bad type non-exclusion condition. If the LP offers a unique contract specifying $k = k_g^*$, then he obtains a payoff of $\nu_g \cdot \pi_g^*$, as he only invests in the fund if it is managed by a good GP. On the other hand, offering a pair of contracts with investment levels $\{k_g^*, k_b^S\}$ 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 $\upsilon_b \cdot \pi_b^S$ from potentially contracting with a bad GP exceeds the expected transfer $\upsilon_g \cdot \Delta\theta \cdot k_b^S$ needed to induce a good GP to choose his own contract. We can write this difference as:

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

approach pooling contracts through a path in which pooling contracts are always dominated.

Observe that $R(k) - R'(k) \cdot k$ is strictly increasing in k which, coupled with the assumptions that $\lim_{k\to 0} R'(k) \cdot k = 0$ and $R(0) = 0$, ensures that $\upsilon_b \cdot \pi_b^S - \upsilon_g \cdot \Delta\theta \cdot k_b^S$ is positive. Nonetheless, although excluding a bad GP is never optimal in a static setting, the optimal two-period contract may entail exclusive contracting with a good GP, as we have seen above.

C Omitted Proofs

Proof of Lemma 1.

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

$$\theta_g \cdot (k_{1g} - k_{1b}) + \delta \cdot \Delta\theta \cdot F(\pi_b^*) \cdot k_b^* \leq x_{1g} - x_{1b} \leq \theta_b \cdot (k_{1g} - k_{1b}).$$

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

Conversely, if condition (3) holds, then the bad type's incentive compatibility constraint does not bind, so that the pair of investment levels k_{1g} and k_{1b} is incentive-compatible.

Now, in any optimal contract the bad LP's participation constraint and the good LP's incentive compatibility constraint must bind. Hence, we have that $x_{1b} = \theta_b \cdot k_{1b}$ and that $x_{1g} = \theta_g \cdot k_{1g} + \Delta\theta \cdot (k_{1b} + \delta \cdot F(\pi_b^*) \cdot k_b^*)$. Hence, it follows that the bad type's incentive compatibility constraint constraint reads:

$$x_{1g} - \theta_b \cdot k_{1g} \geq \Delta\theta \cdot (k_{1b} + \delta \cdot F(\pi_b^*) \cdot k_b^* - k_{1g}).$$

Now, observe that $x_{1g} = \theta_b \cdot k_{1g}$ if and only if equation (3) holds with equality. ■

Proof of Corollary 3.

By staying in a partnership, an LP would obtain a profit of $R(k_i^*) - x_{2i}^D$. On the contrary, paying the default penalty P_i and cashing his outside option would lead to a payoff of $I - P_i$. By the previous lemma, we know that $P_i = \Pi_{2i}^D$. Moreover, observe that $R(k_i^*) - x_{2i}^D = \pi_i^* - \Pi_{2i}^D$, as the LP's second period profits are simply given by the surplus created by the partnership π_i^*, once net rents to the LP Π_{2i}^D has been substracted. Hence, we have that the LP will opt out of the partnership whenever $I_i > \pi_i^*$. ■

Proof of Proposition 4.

(i) This item corresponds to the statement in Corollary 1, proved in Lemma 1.

(ii) Assume that $F(\pi_b^*) > F_b^{CI}$ and $F(\pi_b^*) > F_b^D$. By construction of F_b^D, any contract specifying that the first period investment levels be k_b^S and k_g^*, respectively, would be such that $P_g \geq \Delta\theta \cdot k_g^*$, so that the optimal choice for a bad GP impersonating a good one would be to run the fund for two periods. Hence, we can write the bad GP's profit from mimicking a good GP as:

$$\Pi_b\left(C_g^{ND}\right) = \Delta\theta \cdot \left[\left(k_b^S - k_g^*\right) + \delta \cdot \left(F(\pi_b^*) \cdot k_b^* - F\left(\pi_g^*\right) \cdot k_g^*\right)\right] < 0,$$

the latter inequality following from the facts that $k_b^S < k_b^* < k_g^*$ and that $F(\pi_b^*) < F\left(\pi_g^*\right)$. Hence, the bad GP's incentive compatibility constraint would never bind in this case.

(iii) Assume that $F(\pi_b^*) > F_b^{CI}$ and $F\left(\pi_g^*\right) > F_g^D$. By construction of F_g^D, it follows that $P_g < \Delta\theta \cdot k_g^*$, so that the optimal choice for a bad GP is not to run the fund for two periods. Then, a bad GP's payoff from impersonating a good type is given by:

$$\Pi_b\left(C_g^{ND}\right) = -\Delta\theta \cdot k_g^* + \delta \cdot \left(1 - F\left(\pi_g^*\right)\right) \cdot P_g.$$

As long as $\Pi_b\left(C_g^{ND}\right) \leq 0$, the no-distortions separating menu of contracts is incentive compatible. Hence, in this case, the good GP obtains an informational rent of $k_b^S + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*$. Hence, we can write the bad GP's incentive compatibility constraint as:

$$\Pi_b\left(C_g^{ND}\right) = -\Delta\theta \cdot k_g^* + \left(1 - F\left(\pi_g^*\right)\right) \cdot \left(k_b^S + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*\right).$$

Hence, $\Pi_b\left(C_g^{ND}\right) \leq 0$ if:

$$F\left(\pi_g^*\right) \geq F_g^* \equiv 1 - \frac{k_g^*}{k_b^S + \delta \cdot F\left(\pi_b^*\right) \cdot k_b^*}.$$

■

Proof of Proposition 7.

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

$$\frac{\partial \Pi_L}{\partial k_{1b}}\left(k_{1b}^D\right) = \upsilon_b \cdot \left(R'\left(k_{1b}^D\right) - \theta_b\right) + \upsilon_g \cdot \left(\left(R'\left(k_{1g}^D\right) - \theta_g\right) \cdot \left(1 - F\left(\pi_g^*\right)\right) - \Delta\theta\right).$$

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

$$\upsilon_b \cdot R'\left(k_{1b}^D\right) + \upsilon_g \cdot R'\left(k_{1g}^D\right) \cdot \left(1 - F\left(\pi_g^*\right)\right) = \theta_b - \upsilon_g \cdot \theta_g \cdot F\left(\pi_g^*\right).$$

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

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

which is the condition that determines the distortion for contracts without penalties. We have shown above that $k_{1b}^Z < k_b^S < k_g^* < k_{1g}^Z$. We now show that k_{1b}^D approaches k_b^S as $F\left(\pi_g^*\right)$ increases. First, we have that:

$$\frac{\partial^2 \Pi_L}{\partial \left(k_{1b}\right)^2}\left(k_{1b}^D\right) = \upsilon_b \cdot R''\left(k_{1b}^D\right) + \upsilon_g \cdot R''\left(k_{1g}^D\right) \cdot \left(1 - F\left(\pi_g^*\right)\right) < 0.$$

Also,

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

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

$$\frac{\partial \Pi_L}{\partial k_{1g}}\left(k_{1g}^D\right) = \upsilon_b \cdot \left(R'\left(k_{1b}^D\right) - \theta_b\right) \cdot \frac{1}{1 - F\left(\pi_g^*\right)} + \upsilon_g \cdot \left(R'\left(k_{1g}^D\right) - \theta_g - \Delta\theta \cdot \frac{1}{1 - F\left(\pi_g^*\right)}\right),$$

which leads to

$$\frac{\partial^2 \Pi_L}{\partial \left(k_{1g}\right)^2}\left(k_{1g}^D\right) = \upsilon_b \cdot R''\left(k_{1b}^D\right) \cdot \frac{1}{\left(1 - F\left(\pi_g^*\right)\right)^2} + \upsilon_g \cdot R''\left(k_{1g}^D\right) < 0$$

and

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

where the last inequality follows from the fact that $\upsilon_b \cdot \left(R'\left(k_{1b}^D\right) - \theta_b\right) - \upsilon_g \cdot \Delta\theta < 0$ (observe that $\upsilon_b \cdot \left(R'\left(k_b^S\right) - \theta_b\right) - \upsilon_g \cdot \Delta\theta = 0$ and that $R'\left(k_b^S\right)$ is strictly increasing) and concavity of $R(\cdot)$. ■

Proof of 9.

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

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

With first-period separation, the LP's profits are bounded above by[33]:

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

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

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

which we can write as:

$$\left[\Pi_L^S - \pi_b^*\right] + \delta \cdot \left[F\left(I^P\right) \cdot \left(\Pi_L^* - \Pi_L^S\right) + \left(F\left(\pi_b^*\right) - F\left(I^P\right)\right) \cdot \left(\Pi_L^* - \upsilon_g \cdot \pi_g^*\right) - \upsilon_b \cdot \int_{I^P}^{\pi_b^*} I \cdot dF(I)\right], \tag{13}$$

where

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

stands for the LP's profits if the efficient static investments are implemented and appropriate separation informational rents are paid to a good GP. 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 from the construction of I^P that $\Pi_L^* < \upsilon_g \cdot \pi_g^* + \upsilon_b \cdot I$ for any $I \in \left(I^P, \pi_b^*\right)$, so that $\Pi_L^* < \upsilon_g \cdot \pi_g^* + \frac{1}{F\left(\pi_b^*\right) - F\left(I^P\right)} \upsilon_b \cdot \int_{I^P}^{\pi_b^*} I \cdot dF(I)$. Hence, while the first bracketed term in expression (13) 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 $\Pi_L^* < \Pi_L^S < \upsilon_g \cdot \pi_g^* + \upsilon_b \cdot I$. Consider $\alpha' < \alpha''$, so that $F_{\alpha'}$ first-order stochastically dominates $F_{\alpha''}$. Then, $\Pi_L^* - \Pi_L^S$ carries a relatively lower weight than $\Pi_L^* - \upsilon_g \cdot \pi_g^* + \upsilon_b \cdot I$ in expression (13) under $F_{\alpha'}$ than under $F_{\alpha''}$. Hence, the second bracketed term in expression (13) is smaller (larger

[33] 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.

in absolute value) under $F_{\alpha'}$ than under $F_{\alpha''}$, while the first bracketed term in expression (13) 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 $F_{\alpha''}$. ■