A Theory of Liquidity in Private Equity *

Vincent Maurin† David T. Robinson‡ Per Strömberg§

This draft: December, 2020

Abstract

We develop a model of private equity capturing two critical features of this market: moral hazard for General Partners (GPs) and illiquidity risk for Limited Partners (LPs). The equilibrium fund structure incentivizes GPs with a profit share and compensates LPs with an illiquidity premium. GPs may inefficiently accelerate drawdowns to avoid default by LPs on capital commitments. LPs with higher illiquidity tolerance realize higher returns, leading to return persistence for both funds and LPs. With a secondary market for LP claims, fund persistence decreases, but LP persistence remains. The model can rationalize many empirical findings and offers several new predictions.

Keywords: Private Equity, Liquidity Premium, Secondary Market

JEL Codes: G23, G24, G30, D86

*We thank Naveen Khanna, Giorgia Piacentino, Armin Schwienbacher, Bogdan Stacescu, Paolo Volpin, Guillaume Vuillemey (discussants), Ulf Axelson, Bo Becker, Marcus Opp, Morten Sørensen, and seminar participants at Copenhagen Business School, INSEAD, the Swedish House of Finance, TUM, EIEF, the BGSE Summer Forum, the 6th Cambridge Corporate Finance Theory Symposium, the 15th CSEF-IGIER Symposium on Economics and Institutions, the 2019 LBS Private Equity Symposium, the EFA Meeting in Lisbon, the FTG Meetings in Madrid and Pittsburgh and the FIRS Conference in Savannah for helpful comments. All remaining errors are our own. Per Strömberg would like to thank the NASDAQ Nordic Foundation and the Söderberg Professorship in Economics for financial support. David Robinson would like to thank the Bertil Danielsson Professorship and the Erling Persson Professorship for financial support.

†Stockholm School of Economics. E-mail: vincent.maurin@hhs.se.

‡Duke University, NBER. E-mail: davidr@duke.edu.

§Stockholm School of Economics, CEPR and ECGI. E-mail: per.stromberg@hhs.se.

1 Introduction

Private Equity firms are financial intermediaries that invest in illiquid assets on behalf of outside investors. They commonly raise capital through fixed-life, closed-end funds organized as limited partnerships, in which the General Partners (GPs) – the employees of the private equity firm itself – receive capital pledges from institutional investors, known as Limited Partners (LPs). These capital pledges are not fulfilled at the inception of the partnership but drawn down over time by GPs as investments are identified. The capital is only returned when investments are exited, and GPs typically receive a portion of the net return as a performance fee if the fund return on investment exceeds a benchmark.[1]

These institutional arrangements expose LPs to two related sources of liquidity pressure. First, LPs must ensure that they have liquid funds available to meet GP drawdowns in a timely fashion. Reflecting GPs' concern about LPs' liquidity risk, limited partnership agreements often contain harsh provisions designed to prevent LP default. As we document in this paper, these concerns were most serious during two recent episodes of liquidity stress: the Global Financial Crisis of 2009 and the recent coronavirus pandemic. Second, LPs' fund partnership claims cannot be easily traded. Payoffs to LPs thus only occur when GPs exit the fund investments, which can take several years. The recent development of a secondary market has not eliminated LP liquidity problems because partnership claims often trade at substantial discounts during downturns (Nadauld et al. 2018). Confronted with LPs' liquidity risk, GPs view investors' ability to commit capital for the long term as a key strength.

To study the role that these liquidity considerations play in determining the equilibrium performance and structure of the private equity industry, we propose a simple model of delegated portfolio management based on Holmström and Tirole (1997). LP liquidity concerns interact with GP moral hazard to determine the equilibrium excess return that LPs must earn as a liquidity premium. We find that the interaction between LP liquidity considerations and GP moral hazard can explain the so-called return persistence puzzle in private equity, first documented at the GP-level by Kaplan and Schoar (2005) and at the LP

[1] For more details, see Metrick and Yasuda (2010), Robinson and Sensoy (2013), Litvak (2004), or Huether et al. (2020).

level by Lerner et al. (2007). LPs which are better able to bear illiquidity risk consistently outperform their peers. GPs who can attract these better LPs face fewer investment distortions, leading to stronger performance.[2] More generally, moral hazard of GPs and illiquidity risk for LPs can generate many other empirical regularities in the asset class: the observed rigidity in contract terms (Robinson and Sensoy 2013), the excess returns over liquid assets (Kaplan and Schoar 2005), and the closed-end fund structure itself.

In our model, the commonly observed private equity fund structure emerges as a contractual solution to the agency conflict between LPs and GPs. LPs commit capital for a series of investments rather than on a deal-by-deal basis: This makes it easier to provide incentives for GPs to exert effort. The compensation GPs receive in equilibrium is a function of the overall performance of the fund and resembles the carried interest given to fund managers – as in models of cross-pledging by Diamond (1984) and Laux (2001). This optimal investment structure leaves an unresolved commitment problem for LPs who are exposed to an aggregate liquidity shock. When hit by this shock, they may wish to default on their capital commitment to preserve liquidity for other purposes.

The liquidity risk faced by LPs affects GPs through two distinct channels. First, when LPs face a lower likelihood of a liquidity shock, or such liquidity shocks are less costly, the premium they require for long-term investments goes down. With a lower cost of capital, GPs can raise larger funds. Hence, private equity fundraising is negatively related to subsequent returns, in line with the evidence in Kaplan and Strömberg (2009). Second, to avoid the risk of default by LPs who may experience liquidity shocks, GPs choose to call more capital in the early life of the fund. Early investments act as collateral, ensuring that LPs stand by their commitment to provide capital for later investments. This distortion, however, reduces the incentive benefits from diversification across investments and constrains GPs to raise smaller funds at the expense of total profit.

To account for investors' heterogeneous attitudes toward liquidity risk, we introduce two types of LPs in the model. Good LPs face milder liquidity shocks, which both decreases the

[2]Return persistence at the GP level has also been documented in Harris et al. (2014a), and at the LP level by Dyck and Pomorski (2016) and Cavagnaro et al. (2019).

premium required to invest in private equity and also mitigates their commitment problem. Raising capital from good LPs therefore enables GPs to manage larger, more efficient funds. This creates a distinction between premium capital supplied by good LPs and the overall supply of capital, for a given cost of capital. When premium capital is abundant, only good LPs invest in private equity and expected returns are low. As the demand for capital grows, for instance when GPs' investment opportunities improve, premium capital is relatively scarce. GPs then raise capital also from bad LPs, who are more exposed to liquidity shocks.

A key result in our model is that good LPs earn higher returns in equilibrium than bad LPs when they are both present in the market. This result may seem surprising because good LPs are willing to invest at a lower premium than bad LPs due to their higher tolerance to liquidity risk. Their ability to withstand liquidity shocks, however, allows GPs to run more profitable funds by avoiding inefficient acceleration in drawdowns. In equilibrium, capital supplied by good LPs must command a premium so that GPs are indifferent between offering a high expected return to good LPs and a lower expected return to bad LPs. Hence, we find that some GPs cater only to good LPs who earn a higher return and restrict access to bad LPs (high-return funds are "oversubscribed"). We thus explain the performance persistence for LPs and GPs solely as a function of differences in tolerance to liquidity risk. A straightforward extension of our model allowing GPs to have heterogeneous but observable skills generates assortative matching, which strengthens the persistence results.

The final step of our analysis introduces a secondary market for LP investments. When an aggregate liquidity shock hits the economy, bad LPs sell their partnership interests to good LPs with higher illiquidity tolerance. In our model, the secondary market price for partnership claims is endogenously determined by the amount of liquidity available to good LPs. When this amount of liquidity is low, secondary claims trade at a discount. This discount compensates buyers for providing liquidity, consistent with the empirical findings by Albuquerque et al. (2018) and Nadauld et al. (2018).

The presence of a secondary market alters equilibrium fund size, composition, and excess returns in the primary market via two channels. First, a *liquidity effect* lowers the return required by bad LPs to commit capital to the primary market. These investors benefit from the liquidity provided by good LPs in the secondary market which reduces their required return in the primary market. In addition, the secondary market reduces the default risk of bad LPs, who can now exit through a sale instead of defaulting. Hence, the *liquidity effect* drives down required returns in the primary market and GPs' profit increases because they can raise cheaper capital from bad LPs. The second force is an *opportunity cost effect*: Good LPs now have an incentive to save cash rather than invest in the primary market, in order to profit from secondary market discounts in bad times. This opportunity cost effect increases the expected return required by good LPs to commit capital in the primary market. The combination of these two effects lead to a segmentation of the private equity investor base in equilibrium, with good LPs focusing on the secondary market, leaving bad LPs as capital providers in the primary market.

When a secondary market exists, the incentives for GPs to offer higher expected returns to more liquid LPs are reduced. Our model therefore suggests that the decline in GP performance persistence for buyout funds, documented by Harris et al. (2014a), may be explained by the growth of the secondary market. While this implies that expected return differences across GPs should decrease, our model predicts that performance persistence among LPs should still remain. Good LPs can still earn higher returns by focusing on the secondary market where claims trade at a discount during illiquid periods. In addition, a primary market fund commitment generates a higher expected return when made by a good LP, who, unlike a bad LP, never has to exit at a discount in the secondary market.

The remainder of the paper is structured as follows. Section 2 reviews related theoretical work on private equity. We lay out the model in Section 3, and derive the optimal partnership contract in Section 4. In Section 5 we consider different types of investors, which allows us to deliver the central result of our paper: LPs with higher tolerance to liquidity earn higher returns in equilibrium. In Section 6 we introduce a secondary market and show that return persistence may disappear. We gather the empirical implications of our model in Section 7. Section 8 concludes.

2 Related Literature

Starting from Jensen (1989) and Sahlman (1990), the economic structure of private equity partnerships has been interpreted as the solution to agency problems arising from delegated asset management. Our explanation for the fund structure is based on the idea of cross-pledging benefits, originally due to Diamond (1984). As in Cerasi and Daltung (2000) or Laux (2001), bundling of investments creates "inside equity" which makes it less costly to incentivize the agent. Axelson et al. (2009) also apply this insight to private equity but their analysis focuses on the role for third-party debt financing to mitigate over-investment when project quality is private information. In contrast to them, we consider a moral hazard problem for GPs and we are also able to endogenize fund size and aggregate market activity in equilibrium. More importantly, we study the consequences of illiquidity on expected returns and the role of the secondary market.

A number of papers provide models of the excess return of private over public equity and its implications for portfolio choice. Sørensen et al. (2014) and Giommetti and Sørensen (2019) investigate the illiquidity cost of private equity to investors in dynamic portfolio-choice models. In these papers, the cost of private equity is due to risk-averse LPs becoming exposed to additional uninsurable risk. Phalippou and Westerfield (2014) also solve a dynamic portfolio allocation problem for a risk-averse investor. Similar to Sørensen et al. (2014), the cost of illiquid assets arise from suboptimal diversification, but they add the feature that fund capital calls are stochastic. They also consider the possibility of LP defaults and secondary market sales, taking the discount in the secondary market to be an exogenous parameter. In contrast, we provide an equilibrium model of delegated portfolio management, where we endogenize PE fund and compensation structures as well as equilibrium returns in the primary and secondary markets.

Similar to us, Haddad et al. (2017) explain variation in buyout activity as a result of time-varying risk premia in an agency framework. As in the papers mentioned above, the excess return on private equity compensates risk-averse investors for holding an undiversified portfolio, which in turn is necessary to provide incentives for adding value to the investment.[3] They argue that this compensation increases as the overall market risk premium increases, leading to procyclical fundraising activity. Their model does not distinguish between LPs and GPs as it focuses on the relationship between PE investors and their portfolio companies. Our model differs in that we analyze the frictions between GPs and LPs and characterize the liquidity premium as a compensation to LPs, which can generate return persistence.

Other theoretical explanations for private equity performance persistence have been provided in the literature. Hochberg et al. (2014) and Marquez et al. (2014) argue that persistence can arise from learning about GP skill. In Hochberg et al. (2014), LPs learn the skill of the GPs in which they invest over time, leading to informational holdup when GPs raise their next fund. In Marquez et al. (2014) better portfolio companies want to pair with better GPs. In order to look better to portfolio companies, some GPs have an incentive to limit their fund size, which increases performance due to decreasing returns to scale. This, in turn, leaves LPs with rents and makes funds oversubscribed.[4] In contrast, our model can rationalize GP performance persistence without asymmetric information or differences in skill, as a rent provided to the most liquid LPs for providing capital.

Our explanation for performance persistence is closer to Lerner and Schoar (2004), who also argue that GPs have preference for investors with low costs of illiquidity. In their model, however, investors are uncertain about GPs' skill, and sales of LP claims in the secondary market are interpreted as negative signals of GPs' ability. While they do not derive an optimal fund structure, they argue that GPs endogenously limit trading of Limited Partnership claims to screen for "good" LPs. Our model allows for different funds to be raised in equilibrium offering different returns to their investors. We allow for an active secondary market and derive endogenous discounts to NAV that are not due to information asymmetries. We also show that the introduction of a secondary market can explain why GP performance persistence has weakened over time, while LP performance persistence remains.

[3] Ewens et al. (2013) use a similar mechanism to rationalize the high observed required rates of return that GPs use for evaluating PE investments.

[4] Relatedly, Glode and Green (2011) model the persistence in the returns to hedge fund strategies as a result of learning spillovers.

Finally, a few papers have modeled the secondary market for private equity fund shares. In Bollen and Sensoy (2016), a risk-averse LP allocates funds between public equity, private equity, and risk-free bonds, and has to sell PE assets at a discount if hit by an exogenous liquidity shock. They do not aim to determine primary and secondary market returns in equilibrium, but instead calibrate their model to data to determine whether observed returns and discounts can be rationalized. In our model, secondary market discounts are instead endogenously determined as a result of "cash-in-the-market pricing" when liquidity is scarce. In our model both the supply of liquidity and the long-term asset supply (private equity fund claims) is endogenous, unlike in Allen and Gale (2005), whose framework we build on. Finally, the economic force leading to market segmentation in our paper, where more liquid investors focus on the secondary market in the hope of capturing "fire-sale discounts" is reminiscent of the mechanism in Diamond and Rajan (2011).

3 Model

The model has three periods, $t = 0, 1, 2$. The economy is populated with private equity managers called GPs (General Partners) who have access sequentially to two long-term investments. GPs can raise external financing from investors called LPs (Limited Partners). Two frictions constrain GPs' ability to raise financing. First, GPs are subject to moral hazard and need to be incentivized for investments to be profitable. Second, LPs face liquidity shocks and require a premium to hold long-term assets.

3.1 Agents and Investments

General Partners (GPs)

There is a unit mass of risk-neutral GPs with a cash endowment of A in period 0. GPs do not discount future cash flows. They have access to a linear investment technology in period 0 and in period 1. Both investments mature in period 2. An investment pays off R per unit invested in case of success and 0 in case of failure. Following Holmström and Tirole (1997), the probability of success of an investment depends on an unobservable effort choice by the GP. An investment succeeds with probability p if the GP exerts effort. If the GP shirks, the success probability is q and the GP enjoys a private benefit B per unit invested. The success of an investment is independent of the effort choice for the other investment.

Assumption 1. $pR \geq 1 \geq qR + B$ *(NPV)*

This standard assumption implies that investments have positive NPV only if the GP exerts effort. With a non-contractible effort choice, GPs cannot pledge all the investment cash flows because they would shirk to capture the private benefit. As is standard, we assume that the pledgeable income of an investment is lower than the financing cost.

Assumption 2. $p\left(R - \frac{pB}{p^2-q^2}\right) < 1$ *(Limited pledgeability)*

We will show that the left-hand side in Assumption 2 is the maximum payoff GPs can promise to external investors per unit of investment in a fund.[5] Assumption 2 implies that GPs cannot raise external finance without co-investing their own wealth.

Limited Partners (LPs)

There is a mass K of risk-neutral LPs who consume in period 1 and 2. Each LP is endowed in period 0 with 1 unit of cash she can store. We assume that LP aggregate capital K is large compared to GP wealth:

Assumption 3. $K \geq \frac{A}{1-p\left(R-\frac{pB}{p^2-q^2}\right)} - A$ *(External Capital)*

The left-hand side of the inequality is the capital available to LPs. As we will show, the right-hand side is the maximum amount of financing GPs can raise without shirking.

In period 1, an aggregate liquidity shock hits the economy with probability $\lambda \in (0,1)$. The tolerance of a LP to the liquidity shock is captured by his discount factor $\delta \in (0,1)$ for period 2 consumption when the shock hits. Formally, LPs' preferences are given by

$$u(c_1, c_2) = \begin{cases} c_1 + c_2, & \text{with prob.} \quad 1-\lambda \\ c_1 + \delta c_2 & \text{with prob.} \quad \lambda \end{cases}. \tag{1}$$

[5] With one investment opportunity, the pledgeable income is only $p\left(R - \frac{B}{p-q}\right)$. In our model, as we will show, pledgeable income is higher because GPs have access to two independent investments.

where c_t is period t consumption. There are two types $i \in \{L, H\}$ of LPs with discount factor δ_i such that $\delta_L \leq \delta_H$. The fraction of H-LPs (L-LPs) is denoted μ_H ($\mu_L = 1 - \mu_H$). LP type is public information.

The liquidity shock that LPs face is intended to capture a systemic event that decreases investors' appetite for long-term, illiquid assets, such as a "flight-to-liquidity" episode (see Longstaff 2004). Another example of a shock could be a regulatory change that increases the cost of holding illiquid assets and affects many investors.[6] The heterogeneity among LPs would arise from differences in maturity profile, investment horizon or exposure to regulatory shocks among institutional investors in private equity. The preference for liquidity implies that LPs will require a return premium for long-term illiquid investments compared to cash.

3.2 Partnership Contracts and Markets

Partnership Contracts

In period 0, GPs raise capital from LPs by offering investment partnership contracts. For simplicity, we assume a GP must offer the same contract to all investors of a fund. GPs, however, can still select LPs based on their type, which is public information. As a result, it is without loss of generality that each GP raises capital from one type of LPs only. We thus call i-fund a fund with i-LPs as investors.[7]

Definition 1 (Partnership contract)
For $i \in \{L, H\}$, an i-fund partnership contract is given by the total fund size I_i (including the co-investment A by the GP), the share $x_i \in [0, 1]$ of the fund called by the GP for the period 0 investment and a GP fee schedule $f_i(.)$ per unit of investment, such that

$$\forall\, y \in \{0, R(1-x), Rx, R\}, \qquad 0 \leq f_i(y) \leq y. \tag{2}$$

Given x, let $\mathcal{F}_x^{IC}$ be the set of fee schedules such that the GP exerts effort on both investments.

[6]For example, after the Solvency II regulation was introduced in Europe in 2012, insurance companies were required to hold another EUR40 of balance sheet equity for every EUR100 invested in PE assets, which led many insurers to cut back on their private equity investments (see Fitzpatrick 2011).

[7]As discussed in Section 5, GPs may allow different LP types in a fund if they can offer different contract terms. But our main result is robust: LPs with higher tolerance to liquidity shock can earn higher returns.

The terminology fund is used because a GP makes two investments financed jointly if $x \in (0,1)$. Unlike in our model, GPs and LPs do not contract ex-ante on the share of the fund capital called for each investment. We will show, however, that the optimal value of x lies in an interval, that is, the contract merely specifies investment concentration limits, in line with market practice.[8] The compensation schedule specifies the GP fee per unit of investment for each of the four possible cash flows of the fund in period 2. For instance, cash flow $y = Rx$ corresponds to a success of the first investment (share x) and a failure of the second investment (share $1 - x$). For a given unit cash flow y, the total compensation of the GP is equal to the fee multiplied by the fund size, that is, $f_i(y)I$. The expected (unit) fee to the GP in a i-fund is denoted $\bar{f}_i$. The total expected fee is thus $\bar{f}_i I_i$.

A key variable of interest will be the gross expected return $\bar{R}_{PE,i}$ on PE investment for i-LPs, which can be expressed as a function of the fund contractual features. Provided the GP exerts effort on both investments, which is optimal under Assumption 1, we have

$$\bar{R}_{PE,i} := \frac{(pR - \bar{f}_i)I_i}{I_i - A}. \tag{3}$$

The denominator is the LPs' capital contribution. The numerator is the expected payoff from investment net of the expected fee to the GP. The total expected payoff pRI_i is independent of the investment composition $(x_i, 1 - x_i)$ because investments have the same return. The variable $\bar{R}_{PE,i}$ is thus the expected payoff in period 2 for a dollar invested in a i-fund.

Secondary Market

In Section 6, we will let LPs trade partnership claims in a secondary market after the liquidity shock is realized. A unit secondary claim gives the new owner the cash flow rights attached to 1 unit of capital invested. We let P_i denote the price of a unit secondary claim in a type i fund in the liquidity shock state.[9]

[8] Typical limited partnership clauses restrict funds from investing more than 20% of the commitment in any given deal. See Gompers and Lerner (1996).

[9] This normalization implies that the initial LP makes the capital call before selling the claim. If instead the new LP makes the capital call, the price of the claim would be adjusted by the unit cost $1 - x_i$ of the second capital call to reflect the liability born by the buyer. The economics are unaffected.

Commitment Problem of LPs

The key friction of our model is a commitment problem of investors. In our model, as in practice, LPs only provide a share x of the capital committed in period 0. After a liquidity shock in period 1, LPs may prefer to default on their remaining commitment $1-x$. We set the strongest punishment for LPs as a loss of their entire claim to the fund cash flows if they default.[10] Given the period 2 expected payoff $\bar{R}_{PE,i}$ of a claim in a type i fund and its secondary market price P_i, a LP meets the capital call after a shock if and only if

$$\max\left\{\delta_i \bar{R}_{PE,i}, P_i\right\} \geq 1 - x_i. \tag{4}$$

The left-hand side of (4) is the value of the claim to the LP, equal to $\delta_i \bar{R}_{PE,i}$ if held to maturity after the liquidity shock, or P_i if sold in the secondary market. The LP makes the period 1 capital call if the unit value of the claim exceeds the unit cost $1-x_i$ of providing the remaining capital. As we explain below, GPs seek to avoid LP default, and thus, partnership contracts must satisfy constraint (4). Figure 1 summarizes the sequence of actions.

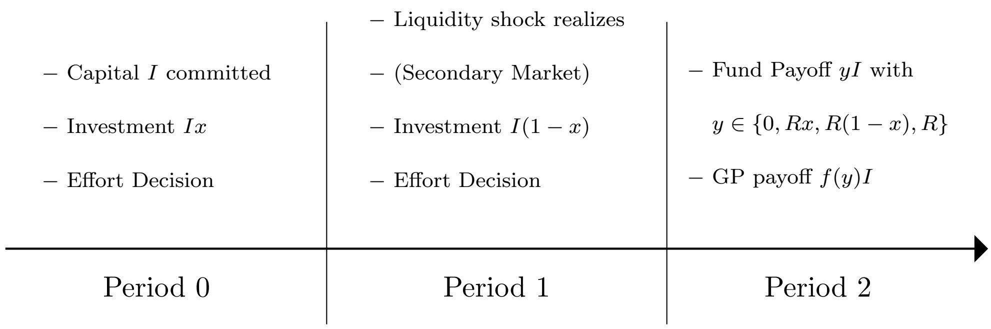

Figure 1: Sequence of actions in a fund. The variables are I: the size of the fund, x: the share of capital invested in period 0, y: the unit fund payoff and $f(y)$: the GP fee per unit of investment.

[10]Quoting from Banat-Estañol et al. (2017), "*default penalties [for LPs] are often written as long lists of punishments, ranging from relatively mild to very severe, implying the loss of some or all of the profits and the forfeiture of the defaulter's entire stake in the fund*". See also Litvak (2004).

Capital Call Waiver?

Partnership contract terms in Definition 1 are not state-contingent. We thus implicitly assume that the liquidity shock is not contractible. GPs, however, may wish to relieve LPs from their capital commitment when the liquidity shock hits since LPs find it costly to provide capital in this state of the world. Such an outcome could be specified as an explicit contractual provision or arise as an ex-post renegotiation outcome between the GP and the LPs.[11] The downside of waiving capital calls for GPs is that they would not be able to carry the second fund investment in full. We capture this trade-off with a non-pecuniary default cost ϕI, proportional to the fund size, incurred by GPs when they cannot invest as planned.

Assumption 4. *The GP default cost verifies* $\phi \geq \max\left\{\frac{pR-1}{\lambda}, \frac{1}{\delta} - pR\right\}$ *(No Default)*

We show in Appendix C.1 that under Assumption 4, GPs find default too costly. Hence, they never waive capital calls for LPs and they design funds such that LPs do not default on capital calls. This result implies that fund design is subject to the no-default constraint of LPs, given by equation (4). To get some intuition about the condition, suppose the GP does not call capital when the liquidity shock hits. Accounting for they expected GP default cost, the NPV of investment per unit is $pR - 1 - \lambda\phi$, which is negative under Assumption 4. Hence, GPs would not raise a fund ex-ante if they expected to default ex-post in period 1.

This default cost assumption is a parsimonious way of capturing several practical elements of partnership arrangements. The deal selection and due diligence that GPs undertake surrounding an investment typically occurs *before* the capital is called from LPs. Our non-pecuniary cost thus proxies for reputational damage and other broken-deal costs if GPs have to back out of deals already agreed upon. Assuming that the reputational damage scales with the foregone profits from the next funds the GP could have raised, it is natural to assume that these costs are proportional to the current fund size.[12] For robustness, we show

[11] There exists anecdotal evidence that some GPs gave capital relief to their LPs during the Great Financial Crisis. In some cases, the LPs who opted in the scheme had to give up part of their claim to the existing investments.

[12] We provide evidence for this interpretation in Section 7. Alternatively, although we use a non-pecuniary cost for simplicity, our assumption can capture in a reduced form operating costs of the fund that GPs usually meet by charging management fees. These fees are proportional to the overall size of capital commitments at the beginning of the fund's life and commonly revert to being proportional to the amount of invested

in Online Appendix D that when the liquidity shock is not too frequent, all our results hold with a pecuniary default cost proportional to the size of the second investment $I(1-x)$.

3.3 Private Equity Equilibrium

We can now define a competitive equilibrium of the model. For a type i-fund, the relevant variables are the expected return $\bar{R}_{PE,i}$ paid by a GP to LPs in period 2 and the price P_i of a secondary claim if there is a secondary market in period 1. An equilibrium is further characterized by the optimal design of each type of fund and the fund choice by GPs. We let $\alpha_i \in [0,1]$ be the fraction of i-funds raised in equilibrium.

Primary Market and Fund Design

In period 0, a GP decides whether to raise a H-fund, a L-fund, or no fund at all, and the terms of the partnership contract in Definition 1. GPs act competitively and take as given the expected return $\bar{R}_{PE,i}$ required by i-LPs for a PE investment. Hence, equation (3) implicitly defines the fund size I_i as a function of the cost of capital $\bar{R}_{PE,i}$ and the expected fee $\bar{f}_i$ in a i-fund. A partnership contract in a i-fund maximizes the profit of the GP if it solves the following problem

$$\max_{x_i, f_i(.) \in \mathcal{F}^{IC}_{x_i}} \Pi_{GP,i} := \bar{f}_i I_i \quad \text{subject to (3) and (4).} \tag{5}$$

The constraint on the fee schedule $f_i(.)$ ensures that the GP exerts effort (see Definition 1) while constraint (4) ensures that LPs do not default on the second investment.

We now characterize the supply of capital from LPs. Let $v_{cash,i}$ (resp. $v_{PE,i}$) be the value to a i-LP of a unit investment in cash (resp. PE) in period 0. As we explain below, the value of cash may be higher than 1 when there is a secondary market for PE claims. For a a PE investment, we have

$$v_{PE,i} := (1-\lambda)\bar{R}_{PE,i} + \lambda \max\left\{\delta_i \bar{R}_{PE,i}, P_i\right\}, \tag{6}$$

capital as the fund matures (see Robinson and Sensoy 2013). Thus, defaults on period 1 capital calls would lower the fee stream that managers receive, introducing disruptions to the normal operations of the fund. This explanation is also consistent with the assumption that the default cost scales with fund size.

The second term of equation (6) reflects the choice for a LP to hold his claim or sell it in the secondary market after a liquidity shock. With linear preferences, a LP invests all his resources in the asset (cash or PE) with the highest value. The supply of capital $S_i(v_{PE,i}, v_{cash,i})$ to type i funds is thus given by

$$S_i(v_{PE,i}, v_{cash,i}) := \begin{cases} 0 & \text{if} \quad v_{PE,i} < v_{cash,i} \\ S \in [0, \mu_i K] & \text{if} \quad v_{PE,i} = v_{cash,i} \\ \mu_i K & \text{if} \quad v_{PE,i} > v_{cash,i} \end{cases} \tag{7}$$

Secondary Market

Finally, we characterize supply and demand in the secondary market for partnership claims. There is no active market for H-fund claims because H-LPs who have the lowest discount rate are the marginal buyers in secondary markets. Hence, in the market for L fund claims, L-LPs are net suppliers of claims while H-LPs are net buyers. The supply from L-LPs is given by:

$$S_L^{sec}(P_L) := \begin{cases} S \in [0, S_L] & \text{if} \quad P_L = \delta_L \bar{R}_{PE,L} \\ S_L & \text{if} \quad P_L > \delta_L \bar{R}_{PE,L} \end{cases} \tag{8}$$

with S_L the primary market commitment. Risk-neutral L-LPs sell their entire participation if the price exceeds their reservation value $\delta_L \bar{R}_{PE,L}$. On the demand side, H-LPs use their resources $\mu_H K$ net of their primary market commitments S_H to buy claims if the price lies below their own higher reservation value $\delta_H \bar{R}_{PE,L}$. Their demand schedule is thus given by

$$D_H^{sec}(P_L) := \begin{cases} \frac{\mu_H K - S_H}{P_L} & \text{if} \quad P_L < \delta_H \bar{R}_{PE,L} \\ D \in \left[0, \frac{\mu_H K - S_H}{P_L}\right] & \text{if} \quad P_L = \delta_H \bar{R}_{PE,L} \end{cases} \tag{9}$$

Definition 2 (Private Equity Market Equilibrium)
An equilibrium is given by a fund composition $\{\alpha_i^*\}_{i=L,H}$*, expected returns* $\{\bar{R}_{PE,i}^*\}_{i=L,H}$*, secondary market prices* $\{P_i^*\}_{i=L,H}$ *and partnership contracts* $(x_i^*, I_i^*, f_i^*(.))$ *such that*

1. *the contract* $(x_i^*, I_i^*, f_i^*(.))$ *solves* (5) *given* $\bar{R}_{PE,i}^*$ *for* $i = L, H$ *(Optimal Partnership),*

2. $\alpha_i^* > 0$ *if and only if* $\Pi_{GP,i}^* = \max\{\Pi_{GP,L}^*, \Pi_{GP,H}^*\}$ *(GP Fund Choice),*

3. $S_i^* = \alpha_i^*(I_i^* - A)$ *(Primary Market clearing),*

4. *the value of cash is* $v_{cash,i}^* = 1$ *and prices satisfy* $P_i^* = \delta_i \bar{R}_{PE,i}^*$*, (No Secondary Market)*

$$\text{or,} \quad v_{cash,i}^* = (1-\lambda) + \lambda \max\left\{1, \delta_i \max_{j=L,H} \frac{\bar{R}_{PE,j}^*}{P_j^*}\right\} \tag{10}$$

and prices satisfy $P_H^* = \delta_H \bar{R}_{PE,H}^*$ *and* $D_H^{sec}(P_L^*) = S_L^{sec}(P_L^*)$*. (Secondary Market)*

The first three optimality criteria for an equilibrium are intuitive. GPs choose the optimal partnership contract for a i-fund taking the cost of capital $\bar{R}_{PE,i}^*$ as given. A type i-fund is chosen in equilibrium if it is the best option among H-funds and L-funds for GPs. Third, the supply of capital from LPs must be equal to the demand from GPs for each type of funds. Without a secondary market, the fourth criterion simply states that LPs derive a marginal utility of 1 from holding cash because cash can only be used for consumption in period 1. With a secondary market, LPs can also buy PE claims with cash in period 1. The value of cash $v_{cash,i}^*$ thus depends on the price of these claims. Finally, secondary markets must clear.

4 Illiquidity Costs of PE

In Section 4 and 5, there is no secondary market for partnership claims. We first characterize the optimal partnership contract for a given type of LPs in Section 4. We thus drop the subscript i for LP type. This analysis sets the stage for Section 5 in which we analyze the private equity market equilibrium and the fund composition with heterogeneous LPs.

Without a secondary market, the return $\bar{R}_{PE}$ required by LPs to invest in private equity takes a simple form. Using equations (6), (7) and the condition $v^*_{cash} = 1$, we obtain

$$\bar{R}_{PE} = 1 + r(\lambda, \delta) := \frac{1}{1 - \lambda(1-\delta)} \tag{11}$$

The net return $r(\lambda, \delta)$ is an illiquidity premium required by investors for long-term investments compared to cash. LPs with higher tolerance to illiquidity δ demand a lower premium to commit capital to private equity. There are gains from trade only if the cost of external capital $\bar{R}_{PE}$ lies below the expected return of investments.

Assumption 5. $pR \geq 1 + r(\lambda, \delta)$ *(Cost of Capital)*

If Assumption 5 did not hold, GPs would only invest their own wealth A. In what follows, we first determine the optimal fee schedule charged by the GP and then the optimal allocation of the fund resources between investments. Part of this analysis relies on the results of Laux (2001): bundling investments into funds arises endogenously because of GP moral hazard considerations.

4.1 Optimal Fee Schedule

GPs face a trade-off between expected fee $\bar{f}$ and fund size I. To see this, consider equation (5) taking as given the return $\bar{R}_{PE}$ required by LPs. Suppose a GP wants to raise a larger fund. Because GP capital A would then represent a lower fraction of the total investment, the expected fee $\bar{f}$ would have to decrease for investors to still earn a return $\bar{R}_{PE}$. The following lemma demonstrates that when facing this trade-off, GPs favor size over fees:

Lemma 1 (Fee Size Trade-off)
GPs minimize the expected fee $\bar{f}$ to maximize fund size I. Problem (5) is equivalent to

$$\min_{x} \min_{f(.) \in \mathcal{F}^{IC}_x} \bar{f} \qquad \textit{subject to} \quad (1-x) \leq \delta(1 + r(\lambda, \delta)) \tag{5b}$$

The constraint faced by GPs is the no-default constraint of LPs, equation (4), in which we replaced $\bar{R}_{PE}$ by its value $1+r(\lambda, \delta)$ thanks to equation (11). To understand why GPs choose size over fees, let us rewrite the GP profit (5) using the implicit capital supply function (3):

$$\Pi_{GP} = pRI - (1 + r(\lambda, \delta))(I - A). \tag{12}$$

The coefficient on I is the expected return on investment net of the external capital cost. Hence, it is optimal to maximize I and thus to minimize $\bar{f}$. We are thus left to determine the value of the first capital call x and the fee schedule that minimize the expected fee $\bar{f}$.

We first derive the optimal fee schedule for a given value of x, denoted $f_x^*(.)$. The schedule specifies a payment to the GP after each of the four possible outcomes for the fund. Under risk-neutrality, it is a well known result that the GP should be paid only after the outcome most informative about effort exertion. Since a success of two independent investments is more informative about effort exertion than a success of a single investment, we have

$$f_x^*(0) = f_x^*(Rx) = f_x^*(R(1-x)) = 0. \tag{13}$$

that is, the GP is paid only if both investments succeed.[13] The incentive-compatibility constraint of the GP then takes a simple form. A fee schedule is in $\mathcal{F}_x^{IC}$ if

$$p^2 f(R) \geq \max\left\{q^2 f(R) + B, pqf(R) + Bx, pqf(R) + B(1-x)\right\}. \tag{14}$$

The three possible payoffs on the right-hand side correspond respectively to the cases in which the GP shirks for both investments, exerts effort only for the second investment and exerts effort only for the first investment. The GP receives private benefits proportional to the fraction of the investment for which he shirks. The probability of a joint success of the investments is reduced to pq (resp. q^2) when shirking on one (resp. two) investments.

By Lemma 1, the GP minimizes the fee charged to LPs. We thus saturate incentive

[13]The reader may observe that when R is small, it can be that $R - f_x^*(R) \leq \max\{Rx, R(1-x)\}$. In this case, the LPs' claim would not be monotonic in the fund cash flows. This monotonicity constraint is often imposed to avoid misreporting of the cash flows by the manager (see for instance Innes 1990). In Internet Appendix E, we solve for the optimal fund design when the monotonicity constraint on the LPs' claim is imposed. We show that the key results from Proposition 1 still hold: Diversification across investments is optimal and these benefits are lower when raising funds from investors with a low value of δ.

constraint (14) to obtain the optimal fee $\bar{f}^*_x$ and the optimal (unconstrained) value of x.

Lemma 2 (Fee Schedule and Diversification)

The optimal expected fee as a function of the share x of capital called in period 0 is

$$\bar{f}^*_x := \begin{cases} \frac{pB}{p-q}(1-x) & \text{if} \quad x \in \left[0, \frac{q}{p+q}\right] \\ \frac{p^2B}{p^2-q^2} & \text{if} \quad x \in \left[\frac{q}{p+q}, \frac{p}{p+q}\right] \\ \frac{pB}{p-q}x & \text{if} \quad x \in \left[\frac{p}{p+q}, 1\right] \end{cases} \tag{15}$$

Hence, the optimal (unconstrained) fraction of capital called at date 0 is $x^* \in \left[\frac{q}{p+q}, \frac{p}{p+q}\right]$.

Lemma 2 shows that the minimum expected fee is U-shaped in the share of capital x called in period 0. This pattern is illustrated on Figure 2. In order to minimize fees, GPs must then distribute the fund capital evenly across investments. Lemma 2 shows that any share x of capital invested in period 0 in the range $\left[\frac{q}{p+q}, \frac{p}{p+q}\right]$ maximizes profit. These diversification benefits arise because GPs can be incentivized more efficiently when two investments are financed jointly. This result is sometimes referred to as *cross-pledging*. As we show below, LPs' default risk may prevent GPs from fully capturing these cross-pledging benefits.

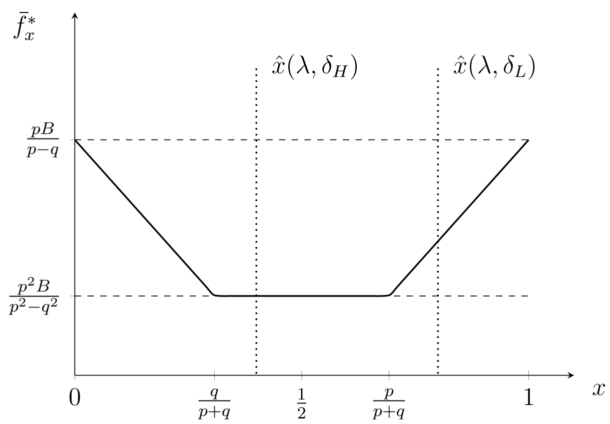

Figure 2: Optimal expected fund fee as a function of the share of capital called for the first investment x. The GP profit is decreasing with the expected fee $\bar{f}$. The vertical dotted lines illustrate the minimum share $\hat{x}(\lambda, \delta)$ a GP can call in period 0 to avoid default by LPs for "liquid" LPs with $\delta_H \geq \hat{\delta}(\lambda)$ and "illiquid" LPs with $\delta_L < \hat{\delta}(\lambda)$.

4.2 Investment Distortion

By Lemma 2, without other constraint, GPs would choose any share x of capital called in period 0 in the range $\left[\frac{q}{p+q}, \frac{p}{p+q}\right]$. As shown by equation (5b), however, the commitment problem of LPs impose a lower bound $\hat{x}(\delta, \lambda)$ on x, with

$$\hat{x}(\lambda, \delta) := 1 - \frac{\delta}{1 - \lambda(1 - \delta)} \tag{16}$$

The LPs commitment problem implies that the first capital call must be large enough to avoid default on the second capital call. Intuitively, the first investment acts as collateral since LPs forfeit the proceeds from the period 0 investment if they default in period 1. When a share x larger than $\hat{x}(\lambda, \delta)$ has been invested, LPs comply with the second capital call. The need to avoid default by LPs can thus lead GPs to distort the optimal investment schedule.

Proposition 1 (Optimal Partnership Contract)
The expected fee $\bar{f}^*_{x^*}$ is given by (15) and the fraction of capital called early is

$$x^* = \begin{cases} x \in \left[\max\left\{\underline{x}, \hat{x}(\lambda, \delta)\right\}, \bar{x}\right] & \text{if} \quad \delta \geq \hat{\delta}(\lambda) \\ \hat{x}(\lambda, \delta) & \text{if} \quad \delta < \hat{\delta}(\lambda) \end{cases} \tag{17}$$

with $\underline{x} = \frac{q}{p+q}$, $\bar{x} = 1 - \bar{x}$, $\hat{x}(\lambda, \delta)$ *is given by equation* (16) *and* $\hat{\delta}(\lambda) := 1 - \frac{p}{p+(1-\lambda)q}$.
The GP inefficiently accelerates capital calls if LPs' tolerance to illiquidity satisfies $\delta \leq \hat{\delta}(\lambda)$.

The key result from Proposition 1 is that to address LPs' default risk, GPs may inefficiently distort the fund investment pattern. GPs need to call and invest enough capital in period 0 to avoid default by their LPs on capital calls in period 1. When liquidity shocks are severe, that is, when $\delta \leq \hat{\delta}(\lambda)$, GPs are forced to call "too much" capital early and do not reap the full incentive benefits from diversification. Equation (16) shows that the lower the tolerance to illiquidity δ, the higher $\hat{x}(\lambda, \delta)$ and thus the more significant the investment distortion. This distortion is costly because it increases the expected fee which reduces fund size and ultimately GP profit. Figure 2 illustrates the results from this section.

The investment distortion faced by GPs can be measured by the difference between the minimum share of capital called early to avoid default, $\hat{x}(\lambda, \delta)$ and the maximum optimal share $\bar{x} = \frac{p}{p+q}$. The threshold $\hat{x}(\lambda, \delta)$ decreases with λ which means that the commitment problem is less severe when a liquidity shock is more likely. When λ increases, the liquidity premium $r(\lambda, \delta)$ paid to LPs also increases and the payoff from staying invested increases, all else equal. The investment distortion is also reduced when p/q increases, a measure of the GP efficiency. When p/q is large, diversification benefits arise for a larger range of values of x, as shown by equation (15). This allows GPs to increase the first capital call x in order to avoid default by LPs while keeping the fee minimal.

It is useful to stress the dual role of liquidity risk in our model. First, investors less sensitive to liquidity risk require a lower liquidity premium $r(\lambda, \delta)$ to invest in PE. This decreases the cost of capital for GPs who respond by raising larger funds in order to increase their profits. Second, due to LPs' default risk, the GP profit is further reduced because he must accelerate investments to avoid LP defaults. As we will see in Section 5, this second feature is essential to explain differences in returns between LPs.

4.3 Robustness

To conclude this section, we discuss alternative ways to mitigate the LP commitment problem, and explain why they do not dominate the fund contract we propose.

To avoid default by LPs, the GP might have to call a significant amount of capital in period 0. Since any capital called is immediately invested in our model, the need to avoid LP default can generate a distortion with respect to the optimal investment schedule. It may seem that the GP could avoid this problem by investing only a fraction of the capital called and hold the rest as cash. We show in Appendix A.1, however, that the GP would then deviate by investing all the capital called in period 0.[14]

Second, the GP could invest proportionally more in period 1 than the LPs to reduce the

[14]One might argue that the GP would be unable to access the excess cash if it is saved in an escrow account until period 1. But LPs might be reluctant to deposit cash with the GP if they have access to a liquid investment opportunity with a strictly positive return. While LPs' outside option is cash for simplicity in our model, it is straightforward to add this feature. Formal derivations are available upon request.

LPs' contribution to the second investment. We show in Appendix **??** that such a scheme would only alleviate the commitment problem of LPs but that it would not solve it. In the main text, for simplicity, and, in line with market practice, we maintain the assumption implicit to Definition 1: Each investor contributes to each investment in proportion to its overall capital contribution.

Alternatively, the GP could "bribe" the LPs when the liquidity shock hits by promising a higher return than in the state with no liquidity shock. We show in Appendix A.3 that this solution is dominated by the benchmark fund structure derived in Proposition 1. The intuition for the result is that the "bribe" is not very valuable to LPs who heavily discount period 2 cash flows after a liquidity shock. As a result, the distortion in the investment schedule is less costly to GPs than the bribe.

Finally, the GP could also raise fresh capital from new investors instead of drawing down on LPs commitments to finance the second investment. In our model, however, the liquidity shock is aggregate and all investors in the economy require a high return on investment when the liquidity shock hits. In fact, we show in Appendix A.4 that the average cost of capital for the GP is strictly higher if he relies on outside capital to finance the second investment.[15]

5 Heterogeneous LPs and Return Persistence

We are now equipped to derive the private equity market equilibrium with two types of LPs. We focus on the interesting case in which the no-default constraint binds for L-LPs.

Assumption 6. $\delta_L < \hat{\delta}(\lambda) < \delta_H$ *(Heterogeneous LPs)*

Assumption 6 implies that H-LPs are better investors than L-LPs. They require a lower rate of return, that is, $r(\lambda, \delta_H) < r(\lambda, \delta_L)$, and their commitment problem is less severe.

[15] This result partially relies on our assumption that the liquidity shock is aggregate. Even if some investors could provide cheaper capital in period 1, there are reasons to believe this solution would be unpractical. The contract would need to specify payouts to GPs and LPs contingent on actions and investments undertaken by third parties, not bound to the initial agreement. These complex contracts would be difficult to implement, not the least because PE funds typically make ten to twenty investments during their investment period. Moreover, raising new capital would divert the GP's attention from ongoing investments. Consistent with these arguments, PE partnership contracts that *require* GPs to raise capital in the future are, to our knowledge, not observed in practice.

5.1 LP Return Persistence

When solving for the private equity market equilibrium, our main variable of interest is μ_H, the share of the total capital owned by H-LPs. When this premium capital is scarce, GPs also raise capital from L-LPs but these worse investors earn a strictly lower return.

Proposition 2 (Equilibrium with Heterogeneous LPs)

There exists thresholds $(\underline{\mu}_H, \bar{\mu}_H)$ *with* $0 < \underline{\mu}_H < \bar{\mu}_H < 1$ *such that*

i) if $\mu_H \leq \underline{\mu}_H$*, both* H*-funds and* L*-funds are raised in period* 0*.* H*-LPs earn a higher return than* L*-LPs,* $\bar{R}^*_{PE,H} > \bar{R}^*_{PE,L} = 1 + r(\lambda, \delta_L)$*.*

ii) if $\mu_H \geq \underline{\mu}_H$*, only* H*-funds are raised in period* 0*, that is,* $\alpha^*_L = 0$*.* H*-LPs earn a net return* $\bar{R}^*_{PE,H}$*, which is decreasing in* μ_H *in the region* $\mu_H \in [\underline{\mu}_H, \bar{\mu}_H]$*. When* $\mu_H \geq \bar{\mu}_H$*,* $\bar{R}^*_{PE,H} = 1 + r(\lambda, \delta_H)$*, that is,* H*-LPs earn their (fair) liquidity premium.*

The expression for the thresholds $\underline{\mu}_H$ and $\bar{\mu}_H$ is provided in the Appendix. Proposition 2 confirms the intuition that GPs do not raise L-funds if capital from H-LPs is abundant (Case ii). When $\mu_H > \bar{\mu}_H$, H-LPs collectively have enough capital to meet the demand from GPs. Because they are on the long side of the market, H-LPs simply earn their liquidity premium to commit capital to a fund. When μ_H is lower than $\bar{\mu}_H$, the resources of H-LPs become scarce and the market clears at an equilibrium rate $\bar{R}^*_{PE,H}$ above the break-even rate of H-LPs. As long as $\mu_H > \underline{\mu}_H$, however, L-LPs cannot compete away these rents because the promised return still falls short of their break-even rate $r(\lambda, \delta_L)$. When the share μ_H of capital available to H-LPs decreases below $\underline{\mu}_H$, the cost of capital for a H-fund becomes so high that some GPs raise funds L-funds.

The key finding of Proposition 2 is that GPs pay a higher return to their investors in a H-fund when both types of funds exist in equilibrium (Case i). Because GPs must distort the investment schedule in a L-fund, L-funds would be strictly less profitable if GPs faced the same cost of capital as for H-funds. Hence, GPs can only be indifferent between funds if the cost of capital is higher for a H-fund, that is, if $\bar{R}^*_{PE,H} > \bar{R}^*_{PE,L}$. This equilibrium premium reflects the higher willingness of GPs to pay for capital supplied by H-LPs. These investors have a lower default risk so that GPs need not call too much capital in period 0.

To be clear, our main finding is not that H-LPs earn a net return over and above their liquidity premium $r(\lambda, \delta_H)$, because this result merely reflects their scarcity. The important result is that H-LPs earn a higher return than L-LPs, due to the commitment problem of investors.[16] We relate this result to the empirical evidence about return persistence for LPs in Section 7. Figure 3 illustrates our findings.

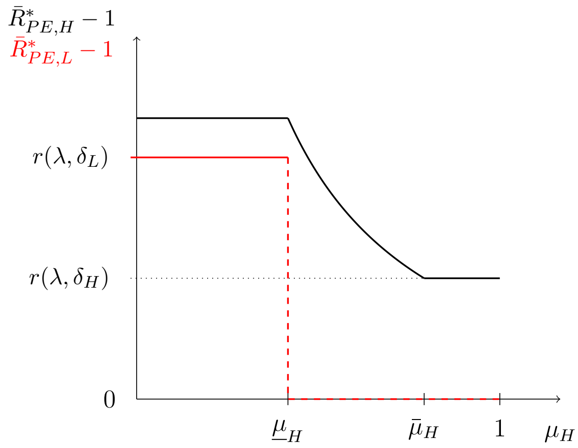

Figure 3: Expected return on PE investment. The variable μ_H is the share of capital owned by H-LPs. There is no L-fund for $\mu_H > \underline{\mu}_H$.

5.2 Return Persistence and Differences in GP skill

The previous section showed that with heterogeneous LPs, different GPs will sometimes offer different expected returns to investors in equilibrium. If we take the liberty of making a dynamic interpretation of our static model, this suggests an explanation for the GP-level performance persistence documented by Kaplan and Schoar (2005). Provided that GPs raising H-funds and L-funds, respectively, tend to stick to their types over time, H-GPs should consistently deliver higher net returns to their LPs. Importantly, the higher expected

[16] If GPs can offer tailored contracts to LPs, funds with a mixed investor composition can arise in equilibrium. Intuitively, if capital supplied by H-LPs is more expensive, GPs would try to raise the minimum amount from H-LPs that avoids the investment distortion, rather than raising capital only from H-LPs. Our key results survive, however: H-LPs earn higher returns and fund segmentation emerges when μ_H is low. Formal results are available upon request.

returns for certain funds does not come from differences in GP skill, but because higher-returning funds cater to more illiquidity-tolerant LPs that are in scarce supply. But what would happen if we did allow for differences in investment skill across GPs? Would more skilled GPs also deliver higher expected returns to LPs?

A simple extension of the model accommodates heterogeneity in GP skills as in Berk and Green (2004). Suppose some GPs are known to have special skills: Their investment pays off R_g in case of success while for other GPs, this payoff is only $R_b < R_g$. Our previous analysis shows that when μ_H is low, GPs compete for scarce capital from H-LPs. GPs win the competition if they are willing to pay a higher rate for this premium capital. Let us denote $\bar{R}^{max}_{PE,H}$ the highest rate a GP is willing to pay for H-LPs' capital. This rate is pinned down by the indifference condition between a H-fund and a L-fund in equilibrium. We have

$$\bar{R}^{max}_{PE,H} = \frac{\bar{f}^*_L(pR - \bar{f}^*_H)}{\frac{\bar{f}^*_H}{1+r(\lambda,\delta_L)}(pR - \bar{f}^*_L) + f^*_L - f^*_H} \tag{18}$$

We show below that $\bar{R}^{max}_{PE,H}$ is strictly monotonic in our proxy for skills, R, which implies that, when GPs have different skills, one type of GP is willing to pay a higher return for H-LPs' capital. If heterogeneity in GP skills are persistent, assortative matching arises at the fund level, but, perhaps surprisingly, H-LPs do not always match with the best GPs.

Corollary 1 (Assortative Matching)
The return $\bar{R}^{max}_{PE,H}$ is increasing in R if and only if $\bar{f}^*_H < 1 + r(\lambda, \delta_L)$. When this condition holds, H-funds are raised by the GPs with higher absolute performance $R_g > R_b$.

Corollary 1 shows that good GPs are not always willing to pay more for premium capital than bad GPs. The intuition is that good GPs also make more profit than bad GPs with L-funds. Hence, positive assortative matching only arises when the relative benefit of avoiding the distortion in L-funds exceeds the profit loss from the higher cost of capital paid to H-LPs. Corollary 1 shows that this is the case when the pledgeable income $pR - \bar{f}^*_H$ in a H-fund is high and the marginal benefit from investment in a L-fund $pR - (1 + r(\lambda, \delta_L))$ is low. In this parameter configuration, when μ_H is low, H-LPs only invest in funds run by skilled GPs because these GPs have a higher willingness to pay for premium capital.

This simple extension explains why the same funds may consistently deliver higher returns to their investors. Our result suggests, however, that positive assortative matching between more skilled GPs and good LPs needs not be the equilibrium outcome. We also stress that return persistence at the fund level would disappear in our model if the only source of heterogeneity is GP skills. Better GPs would simply raise larger funds. Hence, according to our model, differences in LPs liquidity profile is a fundamental source of return persistence while observable heterogeneity in GP skills is not.

6 A Secondary Market for PE Commitments

We now let LPs trade claims in a competitive secondary market in period 1. We will show that the introduction of a secondary market can lead to a radical change in the composition of investors in the primary market. As we explained in Section 3, the secondary market for H-funds is inactive because potential sellers cannot gain from trade with other investors in the economy. Hence, the only relevant market is the secondary market for participation in L-funds. We first take the secondary market price as given and endogenize it in Section 6.2.

6.1 Exogenous Secondary Price

After a liquidity shock hits, L-LPs can now sell their claim. The introduction of a secondary market increases their willingness to commit capital in the primary market. To see this, let us determine the break-even rate $\bar{R}_{PE,L}$ for L-LPs which equalizes the value of cash, $v_{cash,L} = 1$, and the value of private equity $v_{PE,L}$ using equation (6). We have

$$1 = (1 - \lambda)\bar{R}_{PE,L} + \lambda \max\{P_L, \delta_L \bar{R}_{PE,L}\} \quad (19)$$

With a secondary market, L-LPs sell their claim if the price P_L exceeds their reservation value $\delta_L \bar{R}_{PE,L}$. Equation (19) shows that the required return $\bar{R}_{PE,L}$ is strictly lower than the liquidity premium $1 + r(\lambda, \delta_L)$ of L-LPs, defined in (11), when $P_L > \delta_L \bar{R}_{PE,L}$. Since

L-LPs increase their payoff exiting via the secondary market when a liquidity shock hits, they accept a lower nominal return to commit capital in the primary market. We call this effect the *liquidity effect* of a secondary market. In addition, a high secondary market price P_L relaxes the no-default constraint (4) for L-LPs who can sell rather than default. Overall, L-LPs become more suitable investors for PE when partnership claims can be retraded.

The secondary market affects the portfolio choice of H-LPs in an opposite way. As we explained, H-LPs do not use the secondary market to sell their claims. Instead, they act as liquidity providers in the market for L-funds. This opportunity to invest in the secondary market increases the value of cash, rather than the value of a private equity investment. A H-LP is willing to invest in the primary market for PE, if the value $v_{PE,H}$ of a unit investment exceeds the the value of cash $v_{cash,H}$, that is, if

$$\frac{\bar{R}_{PE,H}}{1 + r(\lambda, \delta_H)} \geq (1 - \lambda) + \lambda \max \left\{ 1, \frac{\delta_H \bar{R}_{PE,L}}{P_L} \right\} \tag{20}$$

The value of cash is higher for H-LPs in the presence of a secondary market. When a liquidity shock hits, a dollar buys $1/P_L$ secondary claims to L-funds with unit value $\delta_H \bar{R}_{PE,L}$ to a H-LP. When claims trade strictly below their value for H-LPs, the return on cash is strictly above 1. Equation (20) then shows that the minimum return $\bar{R}_{PE,H}$ required by H-LPs for a PE investment exceeds their break-even rate $1 + r(\lambda, \delta_H)$ due to the *opportunity cost* of committing resources that could have been invested in the secondary market.

The *liquidity effect* for L-LPs is strong when secondary claims trade at a high price P_L while the *opportunity cost effect* for H-LPs is weak. Rather than the price, we use the notion of discount to Net Asset Value (NAV) to capture the liquidity of the secondary market, with

$$D_L := 1 - \frac{P_L}{\bar{R}_{PE,L}} \tag{21}$$

A claim trades at a discount when the secondary market price P_L is lower than the expected value of the claim $\bar{R}_{PE,L}$. A simple argument shows that the equilibrium discount can be no lower than $1 - \delta_H$ and no higher than $1 - \delta_L$. A L-LP would never sell for a price lower than his reservation value, and thus, $P_L \geq \delta_L \bar{R}_{PE,L}$ while a H-LP would only buy if

$P_L \leq \delta_H \bar{R}_{PE,L}$. This result implies, in particular, that claims always trade at a discount following an aggregate liquidity shock. Even investors with the highest valuation for the claim discount period 2 cash flows since $\delta_H < 1$. More interestingly, our analysis will show that the equilibrium discount can increase above this baseline value.[17]

We first state an intermediate result to highlight the strength of the *liquidity effect* and the *opportunity cost effect* as a function of the discount to NAV.

Lemma 3 (Discount to NAV and Returns)
For any discount to NAV, $D \in [1-\delta_H, 1-\delta_L]$, H-LPs require a higher rate of return than L-LPs to invest in the primary market and this difference is increasing in D.

Remember that when there is no secondary market, a result opposite to Lemma 3 holds. Then, the liquidity premium $r(\lambda, \delta_H)$ required by H-LPs lied below the return $r(\lambda, \delta_L)$ required by L-LPs. The presence of a secondary market lowers the minimum rate required by L-LPs through the *liquidity effect*. Simultaneously, the minimum rate required by H-LPs increases because of the *opportunity cost* effect. Lemma 3 shows that these two effects reverse the ranking between required rates.

Lemma 3 implies that, with a secondary market, there is no longer a cost of capital advantage for GPs to raise H-funds. The only rationale left is to avoid the investment distortion arising with L-funds. As we discussed, however, L-LPs are less likely to default when they can resell their claims. Default risk only materializes when the discount to NAV is high but this is also when the cost of capital for H-funds is significantly higher by Lemma 3. This implies that GPs may never prefer to raise H-funds. The Assumption below guarantees this outcome will arise in equilibrium.

Assumption 7. $\left[1 + r(\lambda, \delta_H)\right]\left[1 - \lambda + \frac{\lambda \delta_H (p+(1-\lambda)q)}{(1-\lambda)q}\right] \geq pR$ *(No fund segmentation)*

[17] Our measure of discount can be related to the concepts used in Nadauld et al. (2018). In our model, P is the "return to a seller" while $(1+\bar{R}_{PE})/P$ is the "the return to a buyer". Our economic definition of a discount in (21) prices all investments at their fair value $\bar{R}_{PE,L}$. It is more common in PE to price unrealized investments at cost. The unit value of the second investment would thus be $1-x_L$ instead of $(1-x_L)\bar{R}_{PE,L}$ in definition (21). With this alternative definition, our model is able to generate trades at premium to NAV. Albuquerque et al. (2018) show that PE fund claims sometimes trade at a premium when computed with this alternative method.

As we will show, under Assumption 7, the return H-LPs can make in the secondary market is so high that GPs cannot offer a return acceptable to H-LPs in the primary market.[18]

6.2 Endogenous Discount to NAV

We now endogenize the discount to NAV for L-funds to complete the equilibrium description.

Proposition 3 (Secondary Market Equilibrium)

There are thresholds $0 < \mu_{H,1} < \mu_{H,2} < \mu_{H,3} < \bar{\mu}_H$ for the share of H-LPs capital such that

i) H-funds only exist if $\mu_H \geq \mu_{H,3}$. Then, both LPs earn the same return $\bar{R}^*_{PE,H} = \bar{R}^*_{PE,L}$.

ii) L-funds are the only funds raised for $\mu_H \leq \mu_{H,3}$. The discount to NAV and the share of capital called in period 0 are given respectively by

$$D^*_L = \begin{cases} 1-\delta_L & \text{if} \quad \mu_H \leq \mu_{H,1} \\ \frac{I^*_L - A - \mu_H K}{I^*_L - A - \lambda\mu_H K} & \text{if} \quad \mu_H \in [\mu_{H,1}, \mu_{H,3}] \\ 1-\delta_H & \text{if} \quad \mu_H \geq \mu_{H,3} \end{cases}, \qquad x^*_L = \begin{cases} \hat{x}(\lambda, \delta_L) & \text{if} \quad \mu_H \leq \mu_{H,1} \\ \frac{(1-\lambda)D^*_L}{1-\lambda D^*_L} & \text{if} \quad \mu_H \in [\mu_{H,1}, \mu_{H,2}] \\ \bar{x} & \text{if} \quad \mu_H \geq \mu_{H,2} \end{cases} \tag{22}$$

The expected return in a L fund is given by $\bar{R}^*_{PE,L} = \frac{1}{1-\lambda D^*_L}$.

The key result from Proposition 3 is that H-LPs are not present in the primary market for PE when H-LPs capital is not too abundant, that is when $\mu_H \leq \mu_{H,3}$. Instead of investing in the primary market, H-LPs then focus on the secondary market in which they can buy claims at discount. To understand this result, it is useful to first describe the equilibrium, taking as given that H-LPs only invest in the primary market when $\mu_H \geq \mu_{H,3}$.

Let us consider first the region in which H-LPs capital is abundant. When $\mu_H \geq \mu_{H,3}$, H-LPs and L-LPs are identical investors for GPs since they require the same liquidity premium

[18]We observe that Assumption 7 is only sufficient for our result but imposing it streamlines the analysis. Our objective is to show that the liquidity effect and the opportunity cost effect combined can be so strong as to make H-LPs focus entirely on the secondary market. In Appendix B, we also show, in a limit case, that this result is not generic. When λ is low enough, H-LPs also invest in the primary market in funds that offer higher returns than L-funds. Hence, the primary market segmentation result of Proposition 2 is weaker with a secondary market but it may survive. We discuss the empirical relevance of this result in Section 7.

$r(\lambda, \delta_H)$ to invest in PE. The rate required by L-LPs is lower with a secondary market because they can now sell their claim at a price reflecting the value to an H-LP buyer. When μ_H is large, there are many H-LPs to supply liquidity in the secondary market and L-LPs fully capture the gains from trade. There is a small positive discount to NAV because H-LPs are also affected by the liquidity shock since $\delta_H > 0$. As we observed before, liquidity in the secondary market also mitigates L-LPs default risk, which allows GPs to implement an efficient investment schedule, that is $x_L^* = \bar{x}$.

As H-LPs provide liquidity in the secondary market for L-fund claims, the size of the discount to NAV reflects the scarcity of liquid capital provided by H-LPs, as shown by equation (22). When the share of liquid capital μ_H lies below $\mu_{H,3}$, H-LPs allocate all their resources to the secondary market but they cannot absorb the supply of claims from L-LPs. The market then clears at a cash-in-the-market price: The equilibrium discount now satisfies $D^* > 1 - \delta_H$. In other words, H-LPs receive a compensation for providing liquidity in the secondary market. With lower liquidity in the secondary market, L-LPs require a higher return to invest in the primary market. When secondary claims trade at a low price, L-LPs might now prefer to default on the second capital call. Hence, when $\mu_H \leq \mu_{H,2}$, GPs must distort the investment schedule by setting $x_L^* > \bar{x}$, as in Proposition 2, when there was no secondary market. As shown by equation (22), the discount to NAV for L-funds is capped at $1 - \delta_L$ when H-LPs capital is very scarce. In this case, L-LPs are indifferent between selling and holding their claims and H-LPs capture large discounts in the secondary market.

We can now explain why H-LPs always prefer investing in the secondary market. When there are many H-LPs, they have no comparative advantage with respect to L-LPs in the primary market. Then GPs are not willing to pay more for H capital, unlike in Proposition 2. On the other hand, H-LPs require a higher return because of the *opportunity cost effect* . When μ_H is small, the secondary market is very illiquid and Proposition 3 shows that a L-fund is designed exactly as if there were no secondary market. Then, why do GPs not raise capital from H-LPs to avoid investment distortions in L-funds? The answer is that H-LPs make very large returns from holding cash to buy secondary claims. The *opportunity cost effect* thus disincentivizes investment in the primary market. Under Assumption 7, GPs do not find it profitable to pay the high capital cost required for a H-fund.

6.3 Who gains from a secondary market?

Our discussion showed that secondary trading helps realize a more efficient allocation of capital by creating a missing market between liquid and illiquid LPs when an aggregate shock hits. Building on the results of Proposition 3, we can show that the introduction of a secondary market has redistributive effects between GPs and the most liquid LPs.

Corollary 2 (Welfare effects of a secondary market)
GPs gain from the introduction of a secondary market (strictly if $\mu_H > \mu_{H,1}$). L-LPs are indifferent. H-LPs strictly lose (gain) when their share μ_H of capital is large (small) enough.

For a given return on a PE commitment, L-LPs strictly gain from the introduction of a secondary market because of the *liquidity effect*. But GPs capture these gains from market liquidity by reducing the return promised in the primary market. In equilibrium, L-LPs neither gain nor lose from a secondary market because they are the competitive fringe of this economy. Corollary 2 also shows that GPs gain from the introduction of a secondary market. GPs benefit from the increased competition between investors in the primary market thanks to liquidity provision in the secondary market by H-LPs.

For H-LPs, the welfare effect is ambiguous. When there are few liquid investors, H-LPs earn high returns by pocketing large discounts in the secondary market. By Proposition 6, these returns exceed what they earn with primary PE investments in the absence of a secondary market. However, H-LPs lose when many of their peers are present in the market. By providing liquidity, H-LPs facilitate the competition from L-LPs in the primary market, which lowers their own return. While it is privately optimal for a H-LP to invest in the secondary market, these investors would collectively prefer to shut down the secondary market in this case.

6.4 Does Persistence Persist with Secondary Markets?

We conclude this section by discussing the effect of a secondary market on return persistence, the main result of Section 5. Without a secondary market, we showed in Proposition 2 that GPs promise strictly higher returns to their investors in H-funds when premium capital from H-LPs is scarce enough. Under Assumption 7, this result disappears with a secondary market because only one type of fund is offered in equilibrium. Our model thus suggests that the presence of a secondary market reduces segmentation in the primary market.[19] We can thus attribute part of the recent decrease in fund-level persistence documented by Harris et al. (2014b) to the growth of the secondary market for LP partnership claims.

Even with a secondary market, however, H-LPs still earn higher returns on their PE investment than L-LPs because they invest in a different segment of the market. We showed in Corollary 2 that the average monetary return on a dollar committed to a L-fund for a L-LP is equal to the gross return on cash, equal to 1. By contrast, the average return on a dollar committed to the secondary market by a H-LP is equal to

$$R^*_{cash,H} := \lambda + (1 - \lambda)\frac{R^*_{PE,L}}{P^*_L} > 1 \tag{23}$$

The inequality obtains because $P^*_L < R^*_{PE,L}$ in equilibrium. Hence, H-LPs still earn a higher monetary return than L-LPs by focusing on the secondary market.

We note that an H-LP would also perform better even if constrained to invest only in the primary market. Intuitively, a H-LP would invest and hold his claim while a L-LP sells at a discount when a liquidity shock hits. Because they do not need to rebalance their portfolio in bad times, H-LPs could then achieve higher returns on a PE commitment.

[19] Appendix B shows that under stringent conditions, segmentation can persist. Hence, we do not predict that segmentation should disappear completely.

7 Empirical Relevance

We conclude by relating our model to empirical observations about the PE market. In Section 7.1, we first discuss the relevance of LP default risk in the presence of liquidity funding shocks, the main friction of our model. Section 7.2 reviews the empirical evidence for return persistence, which in our model, arises when LPs are heterogeneous in their tolerance to liquidity risk. Finally, in Section 7.3, we relate our stylized fund structure to actual partnerships and show that our model rationalizes other stylized facts about PE.

7.1 LP Default Risk

The default penalties imposed in PE fund agreements suggest that LP default risk is a real concern of GPs. As already mentioned, Banat-Estañol et al. (2017) and Litvak (2004) find that default penalties, including the forfeiture of the defaulter's stake in the fund, are common (see footnote 10). We further assume that LP defaults are costly to GPs, which we capture with a non-pecuniary default cost. Based on industry commentaries, we believe this is a realistic assumption. As Silveira and Harris (2016) puts it:

> *"[The] GP is exposed to credit risk from any of its LPs having a cash crunch and defaulting. [...] A sudden default can throw off the (usually very tight) deal timetable, exposing the fund to fluctuations in price and cause it to overrun the deal exclusivity period. In the majority of cases, the GP will be forced to borrow money to fund the resulting shortfall, or else face the costs of delay – in an acquisition, this could mean the loss of opportunity and reputation if it fails for insufficient funds."*

Concerns about LP defaults also seem to rise during liquidity crises. Over the last 15 years, financial markets have experienced two episodes that correspond to aggregate liquidity shocks we have in mind in our model: the Great Financial Crisis (GFC) of 2009 and the Covid-19 crisis starting in March 2020. While no systematic data on actual LP defaults exists, there is ample anecdotal evidence that LPs struggled with making capital calls during these crises, and in some cases, defaulted on their commitments.[20] In our model, defaults do not happen in equilibrium because GPs take preventives measures. Actual default events are indeed believed to be rare (Witkowsky 2020) but numerous reports and articles confirm that there is a *risk* of default by LPs during liquidity crises. Among the main reasons cited for these concerns, some institutional investors faced large liquidity outflows that reduced the funds available to meet capital calls.[21]

In our model, GPs respond to LP default risk by accelerating capital calls. While no formal empirical study has yet tested this prediction,[22] several news articles report that GPs accelerated their capital calls in reaction to LP liquidity risk during both the GFC and the Covid-19 crises.[23] During the GFC, Griffith (2009) observed that

> *"some funds have drawn up to 20% of their capital upon closing [...]. It's a way for GPs to make sure their investors have skin in the game from the start."*

Our model also suggests that selling PE claims in secondary markets provide a way out for liquidity-constrained LPs, and that secondary sales will be done at a discount during liquidity shocks when there is a shortage of liquid buyers. Empirical studies of the PE secondary market by Kleymenova et al. (2012), Nadauld et al. (2018), and Albuquerque et al. (2018) provide evidence consistent with this prediction. In particular, they document that discounts to NAV increased dramatically during the GFC, and that the composition of funds traded in the secondary market switched towards younger funds with higher remaining capital commitments. Albuquerque et al. (2018) explicitly relate discounts to liquidity-driven

[20] See e.g Griffith (2009); Zeisberger et al. (2017) p. 292; Lynn (2020); and Bippart (2020).

[21] A survey by placement agency Campbell Lutyens found that "*40 percent of more than 150 LP respondents to its LP sentiment survey published April 4 were not concerned about liquidity, while 31 percent were worried about it. Thirteen percent of respondents were already liquidity constrained*". Another source of concern for LPs is the "denominator effect" according to another survey by LP industry association ILPA, also discussed by Witkowsky (2020). Many LPs exceed their mandated illiquid asset allocations in crisis as liquid asset valuation fall more rapidly than stale valuations for PE assets. The survey finds that 63 percent of respondents were either extremely concerned about exceeding their private equity policy allocations

[22] While Ljungqvist et al. (2020) study the determinants of fund drawdowns both theoretically and empirically, their focus is on the effect of time-varying investment opportunities rather than LP liquidity concerns.

[23] A March 2020 survey of LPs by ILPA found that *"more than half of LPs say capital calls have increased since the onset of Covid-19 as GPs look for headroom in anticipation of an investor liquidity crisis"* (Lynn 2020). Mitchenall (2020) quotes one GP saying that *"they have brought forward to April a capital call they were going to issue in the summer, just to be on the safe side."*

trades, and also show that the composition of secondary buyers switches during liquidity crisis periods (such as the GFC) towards what they refer to as "asset owners" (for instance pension funds and sovereign wealth funds) with deep pockets.[24]

7.2 PE fund returns and performance persistence

Starting with Kaplan and Schoar (2005), a substantial literature has studied the returns to private equity funds. While the asset class has outperformed public equities on average over the last thirty years, the excess performance is countercyclical and particularly high when PE fundraising is low. In addition, Kaplan and Schoar (2005) showed that private equity funds seem to exhibit a significant amount of return persistence: GPs who have outperformed their peers in past funds also tend to outperform in future funds. Persistence in returns has also been documented at the LP level, where certain classes of institutional investors have realized significantly higher returns to their private equity portfolios compared to others.[25] Such performance persistence does not seem to be present among fund managers investing in more liquid assets, such as open-ended mutual funds.[26]

Our model provides a new explanation for the performance persistence puzzle for both LPs and GPs. Unlike competing explanations by Hochberg et al. (2014), Marquez et al. (2014) or Glode and Green (2011), our theory does not require information asymmetry about GP skills. In contrast, our model rationalizes GP performance persistence as a rent provided to the most liquid LPs with GPs "cherry-picking" their LPs. GPs value the commitment ability of LPs which allows them to run more profitable funds. Our story, based on LP screening, can rationalize return persistence for veteran fund mangers, for whom information asymmetry should be less severe. We also explain why some GPs restrict access to their funds, preferring to be oversubscribed rather than increasing fund size.

[24]Similar behavior has also been observed during the Covid-19 crisis. For example, Bucak and Mendonca (2020) write that *"Limited Partners who invested in recently closed private equity funds are looking to offload parts of their commitments in the secondaries market, in anticipation of a liquidity crunch."*

[25]For recent evidence on excess returns, Harris et al. (2014a) and Robinson and Sensoy (2016). Return persistence at the GP level has been documented in Kaplan and Schoar (2005) and Harris et al. (2014a), and at the LP level by Lerner et al. (2007), Dyck and Pomorski (2016) and Cavagnaro et al. (2019).

[26]See e.g. Carhart (1997) and Busse et al. (2010)

Our model also yields a novel prediction relating GP performance persistence to the development of secondary markets for LP fund stakes. When a secondary market exists, GPs need not offer higher expected returns to more liquid LPs. Our model therefore suggests that the decline in GP performance persistence for buyout funds, documented by Harris et al. (2014a), may be explained by the growth of the secondary market. While this implies that expected return differences across GPs should decrease, our model predicts that performance persistence across LPs should still remain. Good LPs still earn higher returns by focusing on the secondary market in which they pick up discounts during illiquid periods. In addition, a fund commitment generates a higher expected return when made by a good LP, who, unlike a bad LP, never has to exit at a discount in the secondary market.

7.3 PE fund structures

We note that the optimal contract in our model shares many features of real-world PE fund agreements. As in Axelson et al. (2009), GPs in our model endogenously choose to raise funds rather than financing deals independently, since cross-pledging mitigates the GP-LP agency problem (similar to Laux (2001)). GPs are rewarded by a profit share if fund payoffs are above a certain level, consistent with the carried interest observed in real-world fund agreements. Cross-pledging also provides an explanation for the concentration limits that are standard in private equity, where buyout fund agreements typically require that the investment in any given deal cannot exceed 20% of the total fund commitment. Our model also predicts that GP co-investment should be an integral part of the LP-GP contract, consistent with actual fund agreements.

The moral hazard model we use, based on Holmström and Tirole (1997), also allows us to relate to empirical findings regarding fundraising and optimal fund size.[27] Our model provides an explanation for why previously successful PE funds do not increase their fee levels to a greater extent. Fees in private equity have been shown to vary remarkably little

[27] Axelson et al. (2009) assume a fixed investment size and can therefore not address these issues. On the other hand, their adverse selection model provides an explanation for the prevalent use of third-party debt financing in (leveraged) buyout deals.

across funds and over time, especially when it comes to the carried interest, where 94% of the PE funds in Robinson and Sensoy (2013) have a carried interest of exactly 20%. Our model indeed shows that GPs would rather increase fund size rather than fees when fundraising conditions improve. These authors also show that variation in carried interest is much stronger for VC funds. A simple extension of our model with fixed scale can explain the difference between buyout and VC funds. Under a scalability constraint, a successful fund manager would increase fees rather than size. We believe VC is less scalable than buyout because manager investing in early-stage start-ups cannot as easily scale up the amount invested in any given company, since start-ups are almost by definition bounded in size.[28]

A dynamic interpretation of our model can rationalize additional empirical results. Kaplan and Schoar (2005) find that PE firms raise the size of their funds when previous fund performance has been relatively strong. In our model, successful GPs will have earned higher carried interest and will therefore have more wealth A to invest in their next fund. This, in turn, increases the amount of capital $I - A$ they can raise from LPs. Finally, the empirical finding in Kaplan and Strömberg (2009) that funds raised during strong fundraising periods have lower returns is a straightforward implication of our model, where GPs raise more capital when the equilibrium compensation for illiquidity required by investors (and thus expected returns) are lower.

There are many interesting features of actual fund agreements to which our model does not speak, such as the level and specific structure of management fee and carry and the determinants of the fund life. While we believe that some of these features could be incorporated by introducing additional elements into the model (e.g. fixed operating costs for running a fund), our main goal is not to provide the ultimate explanation for the PE fund structure. Rather, we want to show how liquidity risk and delegated portfolio management affect PE fundraising and expected returns, while providing a reasonable model of PE funds.

[28]This reading is supported by Metrick and Yasuda (2010) who find that buyout managers build on their prior experience by increasing the size of their funds faster than VC managers do, and conclude that the buyout business is more scalable than the VC business. The limited scalability of VC is also supported in Kaplan and Schoar (2005), who find that the sensitivity of fund size to past performance is significantly stronger in buyout compared to VC.

8 Conclusion

This paper provides a model of delegated investment in private equity funds where investors are subject to liquidity risk. We derive the optimal partnership between GPs and LPs with a fund structure and a compensation contract that resemble actual partnership agreements. Because investors face liquidity risk, there is a pecking order for LPs' capital. GPs prefer to raise capital from LPs who are less sensitive to liquidity risk. These good LPs supply capital at a lower cost and are more likely to stand by their capital commitment. This last feature implies that when high-quality capital is scarce, GPs pay a premium to good LPs. GPs thus cherry-pick their investors for their ability to provide long-term capital. Our results rationalize predictable, persistent differences in performance for both PE fund managers and their investors. With a secondary market for LP claims, good LPs migrate from the primary market to the secondary market. Discounts to NAV arise endogenously when equilibrium liquidity is scarce. Finally, fund-level persistence may disappear with a secondary market while LP-level persistence remains.

Our analysis rests solely on two factors: the agency problem between fund managers and investors and the investors heterogeneous exposure to liquidity shocks. Our model does not exhibit investor irrationality or learning about GP skills but we do not necessarily believe such features are not important in practice. Instead, we provide a benchmark with minimal frictions against which one can judge whether the observed patterns are consistent with agents being symmetrically informed and rational. We believe the stylized structure of our model makes it applicable to delegated portfolio management in other illiquid asset classes, such as infrastructure, private credit, or real estate funds.

Appendix

A Contract Robustness

A.1 Hoarding Capital

We prove our claim in the text that GPs cannot avoid the commitment problem of LPs by calling excess capital in period 0. The key intuition for the result is that GPs would deviate by choosing to invest in period 0 all the capital called. We assume that LPs observe the actual investment before the GP chooses the effort level. In this analysis, we focus on the case $\delta < \hat{\delta}(\lambda)$ since otherwise, Proposition 1 shows that the commitment problem of LPs is moot.

Claim A.1. *If $R \geq \frac{B}{(p-q)^2}$, GPs cannot increase profit by holding LPs' capital as cash.*

As the proof below makes clear, the assumption about R implies that LPs cannot commit to punish the GP by taking control of the fund. It is compatible with Assumptions 1 and 2.

Proof. To avoid default by LPs, the GP must call at least a fraction $\hat{x}(\lambda, \delta)$ of the fund capital in period 0 where $\hat{x}(\lambda, \delta)$ is defined in Proposition 1. Without loss of generality, we assume that the GP calls exactly $\hat{x}(\lambda, \delta)$. We denote by $x_{inv} \in \left[\frac{p}{p+q}, \hat{x}(\lambda, \delta)\right]$ the amount that the GP should invest in period 0 according to the partnership contract. The lower bound on x_{inv} is without loss of generality since Proposition 1 shows that the diversification benefits are maximized for this value. Given the fund size I, the GP expected compensation if he invests according to the contractual schedule is given by

$$\Pi_{GP} = \frac{pBx_{inv}}{p-q} I$$

The GP can deviate by investing all the capital $\hat{x}(\lambda, \delta)I$ he calls in period 0. To punish the GP, the LPs could cancel the second capital call but we guess and verify that the GP would still be compensated to exert effort on the first investment. The minimum expected fee compatible with effort after a deviation is $\hat{\hat{f}} = \frac{pB}{p-q}$. The profit of the GP after the deviation is then equal to

$$\hat{\Pi}_{GP} = \frac{pB}{p-q} \hat{x}(\lambda, \delta) I > \Pi_{GP}$$

which implies that a deviation is optimal. The payoff to LPs from the first investment is

$$\hat{\Pi}_{LP} = (pR\hat{x}(\lambda,\delta) - \hat{\bar{f}})I = p\left(R - \frac{B}{p-q}\right)\hat{x}(\lambda,\delta)I$$

We are left to check that the LPs would not try to punish the LPs following a deviation. Suppose that LPs confiscate the proceeds from the first investment. The GP would react by shirking, leaving a profit equal to $q\hat{x}(\lambda,\delta)RI$ to the LPs. Under the condition stated in Claim A.1, this payoff is lower than $\hat{\Pi}_{LP}$. Hence, LPs cannot commit to punishing the GP after a deviation.

We proved that holding excess capital is not incentive-compatible for the GP. Hence, it is optimal to set $x_{inv} = \hat{x}(\lambda,\delta)$: All capital called in period 0 is invested. This concludes the proof. □

A.2 Non-Proportional Contributions

In this section, we assume that the GP's contribution to the second investment can be higher than its share of the total fund capital.

Intuitively, the limited commitment problem of LPs is most relaxed if the GP invests all his capital A in period 1. Hence, for a fund of size I with a first investment xI, LPs contribute a fraction $\tilde{x} = \frac{xI}{I-A}$ of their committed capital to the first investment. Because the timing of investment contributions does not matter to LPs, their participation constraint is still given by equations (3) and (11). The no-default constraint of LPs, stated in equation (5b) becomes

$$\delta(1 + r(\lambda,\delta)) \geq 1 - \tilde{x} = 1 - x\frac{1 + r(\lambda,\delta)}{pR - \bar{f}^*_x} \tag{A.1}$$

where the inequality follows from the participation constraint of LPs. The equation above shows that non-proportional contributions can indeed relax the no-default constraint because the right-hand side of (A.1) is strictly lower than $1 - x$. As we show below, however, the commitment problem of LPs still puts a lower bound on the fraction of the fund capital in the first investment.

Claim A.2. *If the GP can invest all his wealth A in the second project only, the lower bound for* *the share of the first investment is given by*

$$\hat{x}(\lambda,\delta)=\begin{cases}\frac{\left(pR-\frac{pB}{p-q}\right)\left[1-\delta(1+r(\lambda,\delta))\right]}{1+r(\lambda,\delta)-\left[1+\delta(1+r(\lambda,\delta)\right]\frac{pB}{p-q}} & \textit{if} \quad x\leq\frac{q}{p+q}\\ \frac{\left(pR-\frac{p^2B}{p^2-q^2}\right)\left[1-\delta(1+r(\lambda,\delta))\right]}{1+r(\lambda,\delta)} & \textit{if} \quad x\in\left[\frac{q}{p+q},\frac{p}{p+1}\right]\\ \frac{pR\left[1-\delta(1+r(\lambda,\delta))\right]}{1+r(\lambda,\delta)+\left[1+\delta(1+r(\lambda,\delta)\right]\frac{pB}{p-q}} & \textit{if} \quad x\geq\frac{p}{p+q}\end{cases} \tag{A.2}$$

Proof. The result follows from replacing $\bar{f}_x^*$ in equation (A.1) thanks to equation (15). □

A.3 State Contingent Fees

In this section, we let the GP offer a state-contingent fee which implies a state-contingent return for LPs. We denote $\bar{f}_B$ (resp. $\bar{f}_G$) the expected fee in the Bad (resp. Good) state when the (resp. no) liquidity shock hits. We consider only the case $\delta\leq\hat{\delta}(\lambda)$ since otherwise the LP commitment problem is moot. The default fund structure in Proposition 1 corresponds to $x^*=\hat{x}(\lambda,\delta)$ and $\bar{f}_G^*=\bar{f}_B^*=\frac{pB}{p-q}\hat{x}(\lambda,\delta)$.

Claim A.3. *When $\delta<\hat{\delta}(\lambda)$, it is optimal for the GP to choose $x^*=\hat{x}(\lambda,\delta)$ and non-state-contingent fees $\bar{f}_G^*=\bar{f}_B^*=\frac{pB}{p-q}\hat{x}(\lambda,\delta)$.*

Proof. The objective of the GP if he uses state-contingent fees is to lower the first capital call x below $\hat{x}(\lambda,\delta)$, its value in Proposition 1. The benefit is the reduction of the fee in the bad state. By incentive compatibility, the expected fee is given by $\bar{f}_B=\frac{pB}{p-q}x<\bar{f}_B^*$ if $x<\hat{x}(\lambda,\delta)$. Since the GP's objective is to relax the no-default constraint (4), it is intuitive that this constraint binds at the new chosen value of x, that is,

$$(1-x)(I-A)=\delta\left[pR-\bar{f}_B\right]I$$

We will show that a marginal decrease in x below $\hat{x}(\lambda,\delta)$ strictly lowers the profit of the GP.

We first write the GP profit as function of x for $x\in[\bar{x},\hat{x}(\lambda,\delta)]$. Compared to its baseline value $\bar{f}_G^*$, the expected fee in the good state must also change to satisfy the participation constraint of LPs. We have

$$-(I-A)+(1-\lambda)\left[pR-\bar{f}_G\right]I+\lambda\delta\left[pR-\bar{f}_B\right]I=0$$

Combining the equations above, we obtain the following relationship

$$1 - x = \frac{\delta\big[pR - \bar{f}_B\big]}{(1-\lambda)\big[pR - \bar{f}_G\big] + \lambda\delta\big[pR - \bar{f}_B\big]},$$

which allows us to write the GP profit as a function of x. We have

$$\begin{aligned}\Pi_{GP}(x) &= \frac{(1-\lambda)\bar{f}_G + \lambda\bar{f}_B}{1 - (1-\lambda)\big[pR - \bar{f}_G\big] - \lambda\delta\big[pR - \bar{f}_B\big]}A \\ &= \frac{pR - \frac{\delta}{1-x}\left[pR - \frac{pB}{p-q}x\right] - \lambda(1-\delta)\left[pR - \frac{pB}{p-q}x\right]}{1 - \frac{\delta}{1-x}\left[pR - \frac{pB}{p-q}x\right]}A\end{aligned}$$

We can now show that the GP profit is strictly increasing with x for all $x \in [\bar{x},, \hat{x}(\lambda,\delta)]$. The first order derivative with respect to x is given by

$$\begin{aligned}(\Pi_{GP})'(x) \propto\; & \frac{\delta}{(1-x)^2}\left[pR - \frac{pB}{p-q}\right](pR-1) + \lambda(1-\delta)\frac{pB}{p-q}\left(1 - \frac{\delta}{1-x}\left[pR - \frac{pB}{p-q}x\right]\right) \\ & - \lambda(1-\delta)\frac{\delta}{(1-x)^2}\left[pR - \frac{pB}{p-q}\right]\left[pR - \frac{pB}{p-q}x\right] \\ =\; & \frac{\delta}{(1-x)^2}\left[pR - \frac{pB}{p-q}\right]\left(pR - 1 - \lambda(1-\delta)\left[pR - \frac{pB}{p-q}x\right]\right) \\ & + \lambda(1-\delta)\frac{pB}{p-q}\left(1 - \frac{\delta}{1-x}\left[pR - \frac{pB}{p-q}x\right]\right)\end{aligned}$$

Each term on the right-hand side of this equation is positive. Consider the first term. We have $pR - \frac{pB}{p-q}x \leq 1$ by Assumption 2 because $x \in [\bar{x}, \hat{x}(\lambda,\delta)]$. This implies that

$$pR - 1 - \lambda(1-\delta)\left[pR - \frac{pB}{p-q}x\right] \geq pR - 1 - \lambda(1-\delta) \geq pR - (1 + r(\lambda,\delta)) \geq 0$$

The second to last inequality follows from the definition of $r(\lambda,\delta)$ in equation (11) and the last inequality follows from Assumption 5. For the second term, using the same inequality, we have

$$1 - \frac{\delta}{1-x}\left[pR - \frac{pB}{p-q}x\right] \geq 1 - \frac{\delta}{1-x}$$

The term on the right-hand side of this inequality is strictly decreasing with x and it is strictly positive for $x = \hat{x}(\lambda,\delta)$. Combining these observations, the derivative $(\Pi_{GP})'(x)$ is strictly positive for all $x \in [\bar{x}, \hat{x}(\lambda, \delta)]$. It follows that the optimal choice for the GP is $x^* = \hat{x}(\lambda, \delta)$ and thus, there is no benefit from offering state-contingent fees. This concludes the proof. $\square$

A.4 Raising New Capital

In this section, we let the GP raise capital in period 1 from new investors instead of calling existing LPs capital. We show that such a strategy increases the cost of capital for GPs.

Intuitively, there can only be a benefit of raising external capital after a liquidity shock. By linearity, the GP would call no capital from existing LPs in this case. Hence, in equilibrium, LPs invest $x(I - A)$ in period 0 and $(1 - x)(I - A)$ in period 1 if the liquidity shock does not hit. If the liquidity shock hits, LPs do not make the capital call but the amount $(1 - x)(I - A)$ is raised instead from new investors. The GP invests xA in period 0 and $(1 - x)A$ in period 1.

To prove that raising fresh capital in period 1 is sub-optimal, we compare the cost of capital C in our benchmark fund to the cost of capital $\hat{C}$ with the alternative fund structure, for the GP. For a given fund size I,

$$\hat{C} = (1 + r(\lambda, \delta))x(I - A) + (1 - \lambda)(1 - x)(I - A) + \lambda \frac{1}{\delta}(1 - x)(I - A).$$

The first term corresponds to the initial investment by LPs. The second term corresponds to the second capital call made by LPs when no liquidity shock hits. Since LPs do not discount future cash flows in this case, the cost of $(1 - x)(I - A)$ units of capital for the GP is 1. The last term is the cost of funds from external investors when the liquidity shock hits. The unit cost of new capital is equal to $\frac{1}{\delta}$ since new investors are also affected by the aggregate liquidity shock. In our benchmark fund the total financing cost is equal to

$$C = (1 + r(\lambda, \delta))(I - A) = (1 + r(\lambda, \delta))x(I - A) + \frac{(1 - x)(I - A)}{1 - \lambda + \lambda\delta}$$

The first term of C is equal to the first term of $\hat{C}$. Hence, we are left to show that

$$\frac{1}{1 - \lambda + \lambda\delta} \leq 1 - \lambda + \frac{\lambda}{\delta}$$

Define the random variable $\tilde{\delta}$ which takes value 1 with probability $1 - \lambda$ and δ with probability λ.

The left-hand side of the inequality above is equal to $\left(\mathbb{E}[\tilde{\delta}]\right)^{-1}$ while the right-hand side is equal to $\mathbb{E}\left[\tilde{\delta}^{-1}\right]$. Hence, by Jensen's inequality applied to the convex function $x \mapsto 1/x$, the condition stated above holds. This proves that a GP would strictly increase the total cost of capital when raising external capital from new investors.

B Fund Segmentation with a Secondary Market

We prove the claim in footnote 18 that essentially different H-funds and L-funds may coexist in the primary market when λ is low enough.

Claim B.1. *There exists $\hat{\lambda} > 0$ and $\epsilon > 0$ such that when $\lambda \leq \hat{\lambda}$ and $\delta_H - \delta_L < \epsilon$, both L-funds and H-funds are offered in equilibrium with $\bar{R}^*_{PE,H} > \bar{R}^*_{PE,L}$ for $\mu_H \leq \mu_{H,1}$.*

Proof. We consider the allocation in Proposition 3 and show that it cannot be an equilibrium under some parameter configuration. This implies that H-funds are raised and we then show that H-LPs are promised strictly higher returns.

For the first step, we need to show that under some parameters, the maximum rate GPs are willing to pay for H-funds $\bar{R}^{max}_{PE,H}$ exceeds the minimum rate $\bar{R}^{min}_{PE,H}$ H-LPs are willing to accept. When $\mu_H \leq \mu_{H,1}$, the first capital call in a L-fund is $x^*_L = \hat{x}(\lambda, \delta_L)$ in the conjectured equilibrium. We prove the result by continuity considering the limit $\delta_H \to \hat{\delta}(\lambda)$. Using equation (C.6) in the proof of Lemma 3, we obtain

$$\lim_{\delta_H \to \hat{\delta}(\lambda)} \bar{R}^{min}_{PE,H} = (1 + r(\lambda, \hat{\delta}(\lambda)) \left[1 - \lambda + \lambda \hat{\delta}(\lambda) \frac{1 - \lambda(1 - \hat{x}(\lambda, \delta_L))}{(1 - \lambda)(1 - \hat{x}(\lambda, \delta_L))}\right].$$

Taking further the limit when $\delta_L \to \hat{\delta}(\lambda)$, we get

$$\lim_{\substack{\delta_H \to \hat{\delta}(\lambda) \\ \delta_L \to \hat{\delta}(\lambda)}} \bar{R}^{min}_{PE,H} = 1 + r(\lambda, \hat{\delta}(\lambda)) = \lim_{\delta_L \to \hat{\delta}(\lambda)} \bar{R}^{max}_{PE,H}$$

where the second equality follows from (C.11). To prove the desired result, we will show that for some $\delta_L < \hat{\delta}(\lambda)$, we may have $\bar{R}^{min}_{PE,H} < \bar{R}^{max}_{PE,H}$. Because $\hat{x}(\lambda, \delta_L)$ is decreasing with δ_L, it is enough to show that

$$\left.\frac{\partial \bar{R}^{min}_{PE,H}}{\partial x}\right|_{x=\hat{x}(\lambda,\hat{\delta}(\lambda))} < \left.\frac{\partial \bar{R}^{max}_{PE,H}}{\partial x}\right|_{x=\hat{x}(\lambda,\hat{\delta}(\lambda))} \tag{B.1}$$

We obtain
$$\frac{\partial \bar{R}^{min}_{PE,H}}{\partial x}_{x=\hat{x}(\lambda,\hat{\delta}(\lambda))} = \frac{(1+r(\lambda,\hat{\delta}(\lambda))\lambda\hat{\delta}(\lambda)}{(1-\lambda)\big[1-\hat{x}(\lambda,\hat{\delta}(\lambda))\big]^2} = \frac{\lambda(p+q)}{(1-\lambda)q} \tag{B.2}$$

For the right-hand side of (B.1), using equation (C.11), we have

$$\begin{aligned}
\frac{\partial \bar{R}^{max}_{PE,H}}{\partial x} &= \frac{pB}{p-q}\frac{\partial \bar{R}^{max}_{PE,H}}{\partial \bar{f}^*_x} + \frac{\partial \bar{R}^{max}_{PE,H}}{\partial x} \\
&= \frac{pB}{p-q}\frac{(pR-\bar{f}^*_{\bar{x}})(1-\lambda(1-x))\bar{f}^*_{\bar{x}}\big[pR(1-\lambda)-1+\lambda(1-x)\big]}{\big[(1-\lambda)\bar{f}^*_{\bar{x}}\big(pR-\bar{f}^*_x\big)+(\bar{f}^*_x-\bar{f}^*_{\bar{x}})(1-\lambda(1-x))\big]^2} \\
&\quad + \frac{\lambda(1-\lambda)\bar{f}^*_x(pR-\bar{f}^*_{\bar{x}})\bar{f}^*_{\bar{x}}(pR-\bar{f}^*_x)}{\big[(1-\lambda)\bar{f}^*_{\bar{x}}\big(pR-\bar{f}^*_x\big)+(\bar{f}^*_x-\bar{f}^*_{\bar{x}})(1-\lambda(1-x))\big]^2} \\
&= \frac{pB}{p-q}\bar{f}^*_{\bar{x}}(pR-\bar{f}^*_{\bar{x}})\frac{(1-\lambda(1-x))\big[pR(1-\lambda)-1+\lambda(1-x)\big]+x\lambda(1-\lambda)(pR-\bar{f}^*_x)}{\big[(1-\lambda)\bar{f}^*_{\bar{x}}\big(pR-\bar{f}^*_x\big)+(\bar{f}^*_x-\bar{f}^*_{\bar{x}})(1-\lambda(1-x))\big]^2}
\end{aligned}$$

Setting $x=\hat{x}(\lambda,\hat{\delta}(\lambda))=\bar{x}$ in the expression above, we obtain

$$\frac{\partial \bar{R}^{max}_{PE,H}}{\partial x}_{|x=\hat{x}(\lambda,\hat{\delta}(\lambda))} = \frac{\lambda}{1-\lambda} + \frac{(1-\lambda(1-\bar{x})\left[pR-\frac{1-\lambda(1-\bar{x})}{1-\lambda}\right]}{(1-\lambda)\bar{x}\big(pR-\bar{f}^*_{\bar{x}}\big)} \tag{B.3}$$

Note that the second term of (B.3) does not converge to 0 as $\lambda \to 0$. Hence, comparing (B.3) and (B.2), it follows that the required condition (B.1) holds when δ_H and δ_L are close enough to $\hat{\delta}$ and λ is small enough. This implies that the allocation of Proposition 3 cannot be an equilibrium

The second step is to show that H-funds deliver higher return than L-funds in equilibrium. According to Lemma 3, this is true if the equilibrium discount is strictly higher than $1-\delta_H$. Proposition 3 shows that an equilibrium with a discount $D^*=1-\delta_H$ can only exist if $\mu_H \leq \mu_{H,3}$ where $\mu_{H,3} > \mu_{H,1}$. Hence, under the parameter configuration of Claim B.1, the discount is strictly higher than $1-\delta_H$ This concludes the proof. □

C Proofs

C.1 No capital call relief

We prove the claim following Assumption 4: In order to avoid default costs, GPs neither provide capital call reliefs to their LPs ex-post, nor design funds such that LPs would default in equilibrium. We first need to characterize a fund with GP default after a liquidity shock. We then show that under Assumption 4, GPs do no raise financing ex-ante if they expect to default after the shock.

It can be shown that if the GP expects to default after a liquidity shock, he would invest all his wealth A in period 0. Hence, unlike in the main text, proportional contributions from the GP and LPs to each investment are sub-optimal. Obviously, our result is only stronger if we constrain the GP to make proportional contributions to each investment. This feature implies that the cost of the second investment $(1-x)I$ is born entirely by LPs.

We first characterize the outcome after a liquidity shock. The GP defaults either because LPs default themselves or because LPs and the GP agree to reduce the second capital call. In each case, let $(1-x)I^r$ be the effective capital call by LPs in period 1 with $I^r \in [0, I]$ and $\bar{F}^r$ the new expected payoff of the GP. The superscript r stands for renegotiation, whether explicit or implicit.

We first analyze the explicit renegotiation case in which LPs would not default on the baseline contract. We show that renegotiation does not occur ex-post in this case.To see this, let us derive LPs' payoff without and with renegotiation:

$$\Pi_{LP} = -(1-x)I + \delta(pR - \bar{f})I$$
$$\Pi^r_{LP} = -(1-x)I^r + \delta\left(pR\left[xI + (1-x)I^r\right] - \bar{F}^r\right)$$

LPs agree to the renegotiation if $\Pi^r_{LP} \geq \Pi_{LP}$, that is, if

$$(1-x)(1-\delta pR)(I - I^r) \geq \delta(\bar{F}^r - \bar{f}I)$$

The right-hand side of the inequality is the change in GP profit from renegotiation, abstracting from the default cost. Hence, renegotiation to $I^r < I$ can only benefit both parties if $\delta pR \leq 1$. By linearity, it is optimal to renegotiate to $I^r = 0$. The new expected compensation of the GP is

$$\bar{F}^r = \bar{f}I + \left(\frac{1}{\delta} - pR\right)(1-x)I$$

It can easily be verified that this renegotiation is feasible because the expected revenue from the first investment $pRxI$ exceeds the renegotiated fee $\bar{F}^r$ of the GP. But GPs prefer not to renegotiate because the expected gain $\bar{F}^r - \bar{f}I$ is lower than the default cost ϕI under Assumption 4.

Consider now the case in which LPs would default on the period 1 capital call absent renegotiation, that is, $\Pi_{LP} < 0$. Default by LPs is an implicit renegotiation to $I^r = A$ and $\bar{F}^r = pRxI$. Our analysis above shows that the GP cannot benefit from renegotiation because the LP contribution is already 0 after default. The expected payoff for the GP in such a fund is given by

$$\tilde{\Pi}_{GP} = \Big[(1-\lambda)\bar{f} + \lambda(pRx - \phi)\Big] I$$

with $\bar{f}$ the expected fee, x the share of the fund capital called early and I the total fund size. As in the main text, the fund size can be derived as a function of $\bar{f}$ and x using the participation constraint of LPs. LPs break even on their investment if

$$-(xI - A) + (1-\lambda)\Big[(pR - \bar{f})I - (1-x)I\Big] = 0.$$

Hence, the fund size if given by

$$I = \frac{A}{(1-\lambda)(1 - pR + \bar{f}) + \lambda x} \tag{C.1}$$

Taking the first order derivative of the profit with respect to the expected fee $\bar{f}$, we obtain

$$\frac{\partial \tilde{\Pi}_{GP}}{\partial \bar{f}} = -(1 - \lambda + \lambda x)(pR - 1) + \lambda\phi > 0$$

where the inequality follows directly from Assumption 4. This condition means that a GP charges the maximum expected fee $\bar{f} = pR$ which implies that he does not raise capital from LPs. We show in the main text that under Assumption 5, our benchmark fund without default strictly dominates this autarkic outcome. This concludes the proof.

C.2 Proof of Proposition 1

Building on our analysis in the main text, we are left to show that the GP does not increase the expected return over $r(\lambda, \delta)$ in order to relax the no-default constraint (4). Under the binding no-default constraint (4), the expected return to the LPs would be given by $\bar{R}_{PE}(x) = (1-x)/\delta$ where we highlight the dependence on $x \in [\bar{x}, \hat{x}(\lambda, \delta)]$. The expected GP fee is given by $\bar{f}^*_x = \frac{pBx}{p-q}$ by Lemma 2. The GP profit as a function of x is given by

$$\tilde{\Pi}_{GP}(x) = \frac{A\bar{f}^*_x}{1 - \frac{\delta}{1-x}\left(pR - \bar{f}^*_x\right)}$$

Increasing the promised return over $1 + r(\lambda, \delta)$ is sub-optimal if $\tilde{\Pi}_{GP}(\hat{x}(\delta)) > \tilde{\Pi}_{GP}(\bar{x})$. Since the numerator is increasing in x, it is enough to show that the denominator is decreasing in x. We have

$$\begin{aligned} 0 \leq \frac{\partial \tilde{\Pi}_{GP}}{\partial x} \quad &\Leftarrow \quad 0 \leq -\frac{\delta}{1-x}\frac{pB}{p-q} + \frac{\delta}{(1-x)^2}(pR - \frac{pBx}{p-q}) \\ &\Leftarrow \quad 0 \leq \frac{\delta}{(1-x)^2}\left[pR - \frac{pB}{p-q}\right] \end{aligned}$$

The last inequality follows from Assumption 1. This concludes the proof.

C.3 Proof of Proposition 2

We first prove a preliminary result before analyzing the different cases of Proposition 2.

Lemma C.1

If $\alpha_L^* > 0$, then $\bar{R}^*_{PE,L} = 1 + r(\lambda, \delta_L)$

Proof. Equations (6) and the fact that $v^*_{cash,L} = 1$ imply that $\bar{R}^*_{PE,L} \geq 1 + r(\lambda, \delta_L)$. To prove the reverse inequality, we proceed by contradiction. If $\bar{R}^*_{PE,L} > 1 + r(\lambda, \delta_L)$, L-LPs invest all their resources in PE funds, that is $S^L = \mu_L K$. Let us now prove that H-LPs would also invest all their resources in PE funds. Let $\mu_H > 0$ and suppose $\alpha_H^* = 0$. From equation (7) and by market clearing it must be that

$$\bar{R}^*_{PE,H} \leq 1 + r(\lambda, \delta_H) \leq 1 + r(\lambda, \delta_L)$$

This leads to a contradiction because H-LPs capital would be cheaper and H-funds are more profitable for a given cost of capital since $\bar{f}_H^* < \bar{f}_L^*$ by Assumption 6 and Proposition 1. This implies that $\alpha_H^* > 0$ when $\alpha_L^* > 0$. For both funds to be raised in equilibrium, it must be that

$$\Pi^*_{GP,H} = \bar{f}_H^* I(\bar{R}^*_{PE,H}, \bar{f}_H^*) = \bar{f}_L^* I(\bar{R}^*_{PE,L}, \bar{f}_L^*) \geq \Pi^*_{GP,L} \tag{C.2}$$

This condition implies that $\bar{R}^*_{PE,H} > \bar{R}^*_{PE,L}$ since the GP profit is decreasing with the expected fee. and $\bar{f}_H^* < \bar{f}_L^*$. Hence, since $\bar{R}^*_{PE,L} > 1 + r(\lambda, \delta_L)$ and $\delta_H > \delta_L$, we have $\bar{R}^*_{PE,H} > 1 + r(\lambda, \delta_H)$. Equation (7) then implies that $S^H = \mu_H K$. Hence, the total supply of capital by LPs is K while by assumption 3, the demand from GPs is strictly below K. This cannot be an equilibrium since markets would not clear. Our analysis also establishes that $\bar{R}^*_{PE,H} \in [1 + r(\lambda, \delta_H), \bar{R}^{max}_{PE,H}]$ with

$\bar{R}^{max}_{PE,H}$ the value of $\bar{R}_{PE,H}$ that satisfies (C.2) as an equality. In addition, we must have $\alpha^*_L = 0$ when $\bar{R}^*_{PE,H} < \bar{R}^{max}_{PE,H}$. We now examine the three possible cases. □

Proofs for three possible cases

Case 1 $\bar{R}^*_{PE,H} = 1 + r(\lambda, \delta_H)$. This implies that $\alpha^*_L = 0$ by Lemma C.1. By equation (7), the supply S^*_H of capital from H-LPs lies in $[0, \mu_H K]$. The demand for H capital from GPs is given by equation (3) with $\bar{R}^*_{PE,H} = 1 + r(\lambda, \delta_H)$. Hence the market can clear at this rate if $\mu_H \geq \bar{\mu}_H$ with

$$\bar{\mu}_H K := \frac{pR - \bar{f}^*_H}{1 + r(\lambda, \delta_H) - (pR - \bar{f}^*_H)} A \tag{C.3}$$

This proves the last part of Case ii) in Proposition 2.

Case 2 $\bar{R}^*_{PE,H} \in (1 + r(\lambda, \delta_H), \bar{R}^{max}_{PE,H})$. By Lemma C.1, $\alpha^*_L = 0$. From (7), the supply of capital from H-LPs is given by $\mu_H K$ so market clearing imposes the following relationship

$$\mu_H K = I^*_H - A = \frac{pR - \bar{f}^*_H}{\bar{R}^*_{PE,H} - (pR - \bar{f}^*_H)} A, \tag{C.4}$$

which implicitly defines $\bar{R}^*_{PE,H}$ as as strictly decreasing function of μ_H. Comparing (C.3) and (C.4), the inequality $\bar{R}^*_{PE,H} > 1 + r(\lambda, \delta_H)$ implies that this outcome can only be an equilibrium if $\mu_H \leq \bar{\mu}_H$. Because $\bar{R}^*_{PE,H}$ is decreasing with μ_H, the inequality $\bar{R}^*_{PE,H} < \bar{R}^{max}_{PE,H}$ is satisfied if μ_H is higher than some threshold $\underline{\mu}_H$. This threshold is such that $\Pi^*_{GP,H} = \Pi^*_{GP,L}$ with $\bar{R}^*_{PE,H}$ given by equation (C.4) and $\bar{R}^*_{PE,L} = 1 + r(\lambda, \delta_L)$. The threshold $\underline{\mu}_H$ thus satisfies

$$(\underline{\mu}_H K + A)\bar{f}^*_H = \frac{\bar{f}^*_L A}{1 - \frac{1}{1+r(\lambda,\delta_L)}(pR - \bar{f}^*_L)}$$

Case 3 $\bar{R}^*_{PE,H} = \bar{R}^{max}_{PE,H}$. In this case, the capital supply from L-LPs is indeterminate since $\bar{R}^*_{PE,L} = 1 + r(\lambda, \delta_L)$ by definition of $\bar{R}^{max}_{PE,H}$. The supply of capital from H-LPs is $S_H = \mu_H K$ since $\bar{R}^*_{PE,H} > 1 + r(\lambda, \delta_H)$. Hence, market clearing for H-funds requires

$$\alpha^*_H (I^*_H - A) = \mu_H K$$

which pins down α^*_H since I^*_H is only a function of $\bar{R}^*_{PE,H} = \bar{R}^{max}_{PE,H}$ and $\bar{f}^*_H$. This equation implies that α_H^* is an increasing function of μ_H over $[0, \underline{\mu}_H]$ with $\alpha_H^*(\underline{\mu}_H) = 1$.

Hence, we showed that the allocation in Proposition 2 is the unique equilibrium.

C.4 Proof of Corollary 1

Using equation (18), we obtain

$$\frac{\partial \bar{R}^{max}_{PE,H}(R)}{\partial R} \propto \bar{f}_L^* - \bar{f}_H^* - \frac{\bar{f}_L^* \bar{f}_H^*}{1 + r(\lambda, \delta_L)} + \frac{\bar{f}_H^*}{1 + r(\lambda, \delta_L)} \bar{f}_H^* = (\bar{f}_L^* - \bar{f}_H^*) \left[1 - \frac{\bar{f}_H^*}{1 + r(\lambda, \delta_L)}\right]$$

which proves that $\bar{R}^{max}_{PE,H}(R)$ is either strictly decreasing or strictly increasing with R since $\bar{f}_L^* > \bar{f}_H^*$.

We now prove the matching result when $\bar{f}_H^* < 1 + r(\lambda, \delta_L)$. We need to show that the equilibrium return for H-LPs is strictly above the threshold $\bar{R}^{max}_{PE,H}(R_b)$ when μ_H is low. Suppose by contradiction that $\bar{R}^*_{PE,H} \leq \bar{R}^{max}_{PE,H}(R_b)$. Under the condition above, good GPs strictly prefer to raise funds from H-LPs since $\bar{R}^{max}_{PE,H}(R_g) > \bar{R}^{max}_{PE,H}(R_b)$. Their demand for capital is thus strictly bounded below by 0. But as $\mu_H \to 0$, the supply of capital from H-LPs converges to 0. This cannot be an equilibrium. Hence, when μ_H is too low, it must be that $\bar{R}^*_{PE,H} > \bar{R}^{max}_{PE,H}(R_b)$ to clear the market. When this is the case, H-LPs only supply capital to good GPs.

C.5 Proof of Lemma 3

Define $\bar{R}^{min}_{PE,i}(D_L)$ as the minimum rate a type i-LP accepts to invest in the primary market as a function of the discount to NAV in (21) with $D_L \in [1 - \delta_H, 1 - \delta_L]$. This rate is pinned down by the indifference condition between cash and private equity, that is, $v_{cash,i} = v_{PE,i}$. Using equation (6) and the fact that $v_{cash,L} = 1$ for L-LPs, we obtain

$$1 = \bar{R}^{min}_{PE,i}(D_L)\left[1 - \lambda + \lambda(1 - D_L)\right] \quad \Leftrightarrow \quad \bar{R}^{min}_{PE,L}(D_L) = 1 + \frac{\lambda D_L}{1 - \lambda D_L} \tag{C.5}$$

For H-LPs, using equations (6) and (10), we obtain

$$1 - \lambda + \lambda \delta_H \frac{1}{1 - D_L} = \bar{R}^{min}_{PE,H}(D_L)\left[1 - \lambda + \lambda \delta_H\right] \quad \Leftrightarrow \quad \bar{R}^{min}_{PE,H}(D_L) = 1 + \frac{\lambda \delta_H D_L}{1 - D_L}(1 + r(\lambda, \delta_H)) \tag{C.6}$$

Subtracting these two equations, we obtain

$$\bar{R}^{min}_{PE,H}(D_L)-\bar{R}^{min}_{PE,L}(D_L)=\frac{\lambda D_L}{1-\lambda(1-\delta_H)}\left[\frac{\delta_H}{1-D_L}-\frac{1-\lambda(1-\delta_H)}{1-\lambda D_L}\right]=\frac{(1-\lambda)\lambda D_L}{1-\lambda(1-\delta_H)}\frac{\delta_H-(1-D_L)}{(1-D_L)(1-\lambda D_L)} \tag{C.7}$$

Since the numerator is increasing in D_L and the denominator is decreasing, it follows that $\bar{R}^{min}_{PE,H}(D_L)-\bar{R}^{min}_{PE,L}(D_L)$ is strictly increasing in D_L. This difference is then strictly positive for any value of $D_L > 1-\delta_H$ since it is equal to 0 for $D_L = 1-\delta_H$.

C.6 Proof of Proposition 3

The proof is in two steps. We first characterize the equilibrium under the conjecture that H-LPs do not participate in the primary market when $D_L > 1-\delta_H$ (Step 1). We then verify that GPs find optimal not to raise H-funds (Step 2).

Step 1. Equilibrium characterization

Observe first the participation constraint of L-LPs, equation (7), implies that $1 = v^*_{cash,L} = v^*_{PE,L}$. As we showed in Lemma 3, this implies that $\bar{R}^*_{PE,L} = 1+\frac{\lambda D^*_L}{1-\lambda D^*_L}$ as stated in Proposition 3. Second, the period 0 capital call in a L-fund is either the optimal value $\bar{x}$ or such that the no-default constraint binds that is $x^*_L = 1-P^*_L$ which proves the second equality in (22). We are left to determine the equilibrium value of the discount to NAV D^*_L. The proof is by construction. For each possible value of the discount D^*_L, we characterize the range of values for μ_H where this equilibrium exists.

Case 1. $D^*_L = 1-\delta_H$.

From Lemma 3, we obtain that $\bar{R}^{min}_{PE,L}(D_L) = \bar{R}^{min}_{PE,H}(D_L) = 1 + r(\lambda,\delta_H)$ and $x^*_L = \bar{x}$ since $\bar{x} \geq \hat{x}(\lambda,\delta_H)$ by Assumption 6. Since the investment schedule is not distorted in L-funds, by optimality of the fund choice for GPs, it must be that $\bar{R}^*_{PE,L} = \bar{R}^*_{PE,H}$. By the clearing condition in the primary market, we further obtain that $\bar{R}^*_{PE,L} = 1 + r(\lambda,\delta_H)$ since otherwise the supply of funds from LPs would exceed the demand by GPs. We are now left to derive the conditions such that the conjecture $D^*_L = 1-\delta_H$ is satisfied. It must be that the supply of claims in the secondary market exceeds the supply at price $P^*_L = \delta_H(1 + r(\lambda,\delta_H))$, that is,

$$\left[I^*_H - A - S^*_H\right]\delta_H\left[1 + r(\lambda,\delta_H)\right] \leq \mu_H K - S^*_H$$

Note that because $\delta_H\left[1 + r(\lambda,\delta_H)\right] < 1$, this inequality is easier to satisfy when $S^*_H = 0$. Hence, this allocation is an equilibrium for $\mu_H \geq \mu_{H,3}$ with $\mu_{H,3}$ the minimum value of μ_H such that the equation above when setting $S_H = 0$, that is,

$$\mu_{H,3} = \frac{\left[I_H^* - A\right]\delta_H\left[1 + r(\lambda, \delta_H)\right]}{K} \tag{C.8}$$

with I_H^* the fund size given by equation (3) with $\bar{R}_{PE,H}^* = 1 + r(\lambda, \delta_H)$ and $\bar{f}_H^*$ the return and expected fee, respectively. Because $\delta_H\left[1+r(\lambda, \delta_H)\right] < 1$, a comparison between the equation above and (C.3) shows that $\mu_{H,3} < \bar{\mu}_H$.

Case 2: $D_L^* = 1 - \delta_L$.

This implies that $\bar{R}_{PE,L}^{min}(D_L) = 1 + r(\lambda, \delta_L)$. Since by construction GPs only raise L-funds, the clearing condition in the primary market implies that $\bar{R}_{PE,L}^* = 1 + r(\lambda, \delta_L)$. Combining this result with $D_L^* = 1 - \delta_L$, we obtain $x_L^* = \hat{x}(\lambda, \delta_L)$. By equation (8), the supply of claims in the secondary market is indeterminate. This outcome is an equilibrium if the maximum supply of claims in the secondary market exceeds the demand at price $P_L^* = \delta_L(1 + r(\lambda, \delta_L))$. Using equation (9), the condition writes

$$\left[I_L^* - A\right]\delta_L(1 + r(\lambda, \delta_L)) \geq \mu_H K \tag{C.9}$$

with I_L^* the fund size given by equation (3) with $\bar{R}_{PE,L}^* = 1 + r(\lambda, \delta_L)$ and expected fee $\bar{f}_L^*$. We can rewrite this condition as $\mu_H \leq \mu_{H,1}$. Because $I_L^* < I_H^*$, the comparison between equations (C.8) and (C.9) shows that $\mu_{H,1} < \mu_{H,3}$.

Case 3: $D_L^* \in (1 - \delta_H, 1 - \delta_L)$.

Since $P_L^* > \delta_L(1 + \bar{R}_{PE,L}^*)$, L-LPs strictly prefer to sell their claims in the secondary market by equation (8). Because $P_L^* < \delta_H(1 + \bar{R}_{PE,L}^*)$, the liquidity provided by H-LPs is equal to their total resources $\mu_H K$, which they invest only as cash under our maintained conjecture. Hence, the market clearing condition on the secondary market writes

$$\left[I_L^* - A\right]P_L^* = \mu_H K \tag{C.10}$$

By equation (3), the fund size I_L^* is a decreasing function of the expected return $\bar{R}_{PE,L}^* = 1 + \frac{\lambda D_L^*}{1-\lambda D_L^*}$ and of the share of capital called in period 0, $x_L^* = \max\left\{\bar{x}, \frac{(1-\lambda)D_L^*}{1-\lambda D_L^*}\right\}$ which are both increasing function of the discount D_L^*. The price of a claim can be expressed as a (decreasing) function of the discount with $P_L^* = \frac{1-D_L^*}{1-\lambda D_L^*}$. Hence, the term on the left-hand side of (C.10) is a strictly decreasing function of D^*_L. Hence, condition $D^*_L \in (1-\delta_H, 1-\delta_L)$ defines a range of values of μ_H where equation (C.10) may hold. Comparing equations (C.10) to equations (C.8) and (C.9), the upper bound and lower bound of this region are given respectively by $\mu_{H,3}$ and $\mu_{H,1}$.

To complete the equilibrium characterization, let us define $\mu_{H,2}$ as the smallest value of μ_H in $[\mu_{H,1}, \mu_{H,3}]$ such that $x^*_L = \bar{x}$. This threshold is well defined because the function $D^*_L \rightarrow \frac{(1-\lambda)D^*_L}{1-\lambda D^*_L}$ is increasing and strictly below (resp. above) $\bar{x}$ for $D^*_L = 1-\delta_H$ (resp. $D^*_L = 1-\delta_L$).

Step 2. No H-fund when $\mu_H \leq \mu_{H,3}$

We now verify the initial conjecture about H-funds. To prove this result, we show that the minimum return $\bar{R}^{min}_{PE,H}(D^*_L)$ required by H-LPs exceeds the maximum return $\bar{R}^{max}_{PE,H}(D^*_L)$ that GPs are willing to pay. The value of $\bar{R}^{max}_{PE,H}(D^*_L)$ can be derived from equation (18), replacing $1+r(\lambda,\delta_L)$ with $\bar{R}^*_{PE,L}$:

$$\bar{R}^{max}_{PE,H}(D^*_L) = \frac{\bar{f}^*_{x^*_L}\left(pR - \bar{f}^*_{\bar{x}}\right)}{\bar{f}^*_{\bar{x}}\left(pR - \bar{f}^*_{x^*_L}\right) + (\bar{f}^*_{x^*_L} - \bar{f}^*_{\bar{x}})(1+\bar{R}^*_{PE,L})}(1+\bar{R}^*_{PE,L}). \tag{C.11}$$

For $\mu_H \in [\mu_{H,2}, \mu_{H,3}]$, we showed that $x^*_L = \bar{x}$. Hence, equation (C.11) becomes $\bar{R}^{max}_{PE,H}(D^*_L) = \bar{R}^*_{PE,L}$. Since $D^* > 1-\delta_H$, the minimum rate required by H-LPs satisfies $\bar{R}^{min}_{PE,H}(D^*_L) > \bar{R}^*_{PE,L}$ by Lemma 3. This proves our claim in this case.

For $\mu_H \leq \mu_{H,2}$, $\bar{R}^{max}_{PE,H}(D^*_L)$ is increasing with D^*_L. To see this, observe from equation (C.11) that $\bar{R}^{max}_{PE,H}(D^*_L)$ is increasing with x^*_L and $\bar{R}^*_{PE,L}$ which are themselves increasing functions of D^*_L, as shown above. This implies that

$$\bar{R}^{max}_{PE,H}(D^*_L) \leq \bar{R}^{max}_{PE,H}(1-\delta_L) < pR$$

Similarly, we showed in Lemma 3 that $\bar{R}^{min}_{PE,H}(D^*_L)$ is an increasing function of D^*_L. This implies that for $\mu_H \in [\mu_{H,2}, \mu_{H,3}]$,

$$\begin{aligned}\bar{R}^{min}_{PE,H}(D^*_L) \geq \bar{R}^{min}_{PE,H}(D^*_L(\mu_{H,2})) &= 1 + \frac{\lambda\delta_H D^*_L(\mu_{H,2})}{1-D^*_L(\mu_{H,2})}(1+r(\lambda,\delta_H)) \\ &= 1 + \frac{\lambda\delta_H x^*_L(\mu_{H,2})}{(1-\lambda)(1-x^*_L(\mu_{H,2}))\big[1-\lambda(1-\delta_H)\big]} \\ &= 1 + \frac{\lambda\delta_H\bar{x}}{(1-\lambda)(1-\bar{x})\big[1-\lambda(1-\delta_H)\big]}\end{aligned}$$

The result that $\bar{R}^{min}_{PE,H}(D^*_L) > \bar{R}^{max}_{PE,H}(D^*_L)$ obtains because the lower bound on $\bar{R}^{min}_{PE,H}(D^*_L)$ is higher than the upper bound for $\bar{R}^{max}_{PE,H}(D^*_L)$ under Assumption 7. This concludes the proof.

C.7 Proof of Corollary 2

L-LPs

The result for L-LPs is straightforward With or without a secondary market, their value $v_{PE,L}$ from investing in PE is equal to their value for cash $v_{cash,L}$ equal to 1 because they never buy claims in the secondary market. This implies that the utility of L-LPs is always equal to 1, the value of the unit of cash they initially hold.

H-LPs

Let us consider first the case in which μ_H is small. In Proposition 2, we characterized a threshold $\bar{\mu}_H$ below which H-LPs earn strictly higher return than their liquidity premium $r(\lambda, \delta_H)$ on their PE investment. With a secondary market, Proposition 3 showed that H-LPs earn $\bar{R}^*_{PE,H} = 1+r(\lambda, \delta_H)$ for any $\mu_H \geq \mu_{H,3}$. In the proof of Proposition 3, we showed that $\mu_{H,3} < \bar{\mu}_H$. This shows that when $\mu_H \in [\mu_{H,3}, \bar{\mu}_H]$, H-LPs strictly lose from the introduction of a secondary market.

Let us now consider the parameter region $[0, \min\{\mu_{H,1}, \underline{\mu}_H\}]$. By Proposition 2, the return enjoyed by H-LPs without a secondary market is equal to $\bar{R}^{max}_{PE,H}(1-\delta_L)$, which we previously defined as the return that makes a GP indifferent between raising a H-fund or a L-fund. With a secondary market, however, we showed in Proposition 3, that for all values of $D_L \in [1-\delta_H, 1-\delta_L]$, we have $\bar{R}^{max}_{PE,H}(D_L) > \bar{R}^{min}_{PE,H}(D_L)$, which means that H-LPs strictly prefer holding cash than investing in a fund that would deliver return $\bar{R}^{max}_{PE,H}(1-\delta_L)$. This implies that H-LPs are strictly better off with a secondary market.

GPs

By equation (3) and Proposition 1, the profit of a GP and the fund size depends on two variables: the cost of capital $\bar{R}_{PE}$ and the optimal share of the fund capital x^* called in period 0. In this proof, these variables are used as arguments. Without a secondary market, a GP always weakly prefers raising a H-fund with $x^*_H = \bar{x}$ and $R^{prim}_{PE} := R^*_{PE,H}$ given in Proposition 2. With a secondary market, the GP always weakly prefers to raise a L-fund with x^*_L and $R^{sec}_{PE} := R^*_{PE,L}$ given in Proposition 3.

We showed that $\mu_{H,3} < \bar{\mu}_H$ in the proof of Proposition 3. By Proposition 3, for $\mu_H > \mu_{H,3}$, we have $R_{PE}^{sec} = 1 + r(\lambda, \delta_H) \geq R_{PE}^{prim}$ and $x_L^* = \bar{x}$. This implies that GPs' profit is weakly higher for $\mu_H \geq \mu_{H,3}$ and strictly higher for $\mu_H \in [\mu_{H,3}, \bar{\mu}_H]$.

We now consider the parameter region $\mu_H \leq \mu_{H,3}$. We first show that $\mu_{H,1} \leq \underline{\mu}_H$. In the proof of Proposition 3, we showed that

$$\mu_{H,1} K = \delta_L(1 + r(\lambda, \delta_L)) \left[I^*\left(\hat{x}(\lambda, \delta_L), 1 + r(\lambda, \delta_L)\right) - A \right] \leq \left[I^*\left(\hat{x}(\lambda, \delta_L), 1 + r(\lambda, \delta_L)\right) - A \right]$$

In the proof of Proposition 2, we showed that

$$\underline{\mu}_H K = \left[I^*\left(\bar{x}, R_{PE}^{prim}(\underline{\mu}_H)\right) - A \right]$$

in which $R_{PE}^{prim}(\underline{\mu}_H)$ is by definition the return that makes a GP indifferent between a H fund and a L-fund so that

$$I^*\left(\bar{x}, R_{PE}^{prim}(\underline{\mu}_H)\right) = \frac{\bar{f}^*_{\hat{x}(\lambda, \delta_L)}}{\bar{f}^*_{\bar{x}}} I^*\left(\hat{x}(\lambda, \delta_L), 1 + r(\lambda, \delta_L)\right) > I^*\left(\hat{x}(\lambda, \delta_L), 1 + r(\lambda, \delta_L)\right)$$

where the inequality follows from Proposition 1. This proves that $\mu_{H,1} \leq \underline{\mu}_H$. As a result, the profit of a GP is the same with or without a secondary market if $\mu_H \leq \mu_{H,1}$ and it is strictly higher with a secondary market if $\mu_H \in [\mu_{H,1}, \underline{\mu}_H]$ because the cost of capital R_{PE}^{sec} is strictly decreasing with μ_H for all $\mu_H \geq \mu_{H,1}$.

We are thus left to show that GPs earn a higher profit with a secondary market in the region $[\underline{\mu}_H, \mu_{H,3}]$. In this region, using the results in the proof of Proposition 2 and 3 respectively, we have

$$\Pi_{GP}^{prim} = \bar{f}^*_{\bar{x}} I(\bar{x}, R_{PE}^{prim}) = \bar{f}^*_{\bar{x}} [A + \mu_H K]$$
$$\Pi_{GP}^{sec} = \bar{f}^*_{x_L^*} I(x_L^*, R_{PE}^{prim}) = \bar{f}^*_{x_L^*} \left[A + \frac{\mu_H K}{P_L^*} \right]$$

with $P_L^* = \frac{1 - D_L^*}{1 - \lambda D_L^*} < 1$. Since $x_L^* \geq \bar{x}$ by Proposition 3, we obtain $\Pi_{GP}^{prim} < \Pi_{GP}^{sec}$ for $\mu_H \in [\underline{\mu}_H, \mu_{H,3}]$. This concludes the proof.

References

Albuquerque, Rui A., Johan Cassel, Ludovic Phalipou, and Enrique J. Schroth, 2018, Liquidity provision in the secondary market for private equity fund stakes, Working paper, Cass Business School.

Allen, Franklin, and Douglas Gale, 2005, From cash-in-the-market pricing to financial fragility, *Journal of the European Economic Association* 3, 535–546.

Axelson, Ulf, Per Strömberg, and Michael S. Weisbach, 2009, Why are buyouts levered? the financial structure of private equity funds, *The Journal of Finance* 64, 1549–1582.

Banat-Estañol, Albert, Filippo Ippolito, and Sergio Vicente, 2017, Default penalties in private equity partnerships, Working paper.

Berk, Jonathan, and Richard Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Bippart, Graham, 2020, Ilpa to recommend boosting disclosure of capital call credit line usage, https://www.buyoutsinsider.com/ilpa-to-recommend-boosting-disclosure-of-capital-call-credit-line-usage/.

Bollen, Nicolas, and Berk Sensoy, 2016, How much for a haircut? illiquidity, secondary markets, and the value of private equity, Working paper, Fisher College of Business, Ohio State University.

Bucak, Selin, and Elisangela Mendonca, 2020, Liquidity crunch pushing investors to reconsider commitments in secondaries, https://www.fnlondon.com/articles/liquidity-crunch-pushing-investors-to-reconsider-commitments-in-secondaries-20200420

Busse, Jeffrey, Amit Goyal, and Sunil Wahal, 2010, Performance and persistence in institutional investment management, *Journal of Finance* 65, 765–790.

Carhart, Mark, 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Cavagnaro, Daniel R., Berk A. Sensoy, Yingdi Wang, and Michael S. Weisbach, 2019, Measuring Institutional Investors Skill at Making Private Equity Investments, *Journal of Finance* 74, 3089–3134.

Cerasi, Vittoria, and Sonja Daltung, 2000, The optimal size of a bank: Costs and benefits of diversification, *European Economic Review* 44, 1701 – 1726.

Diamond, Douglas, and Raghuram Rajan, 2011, Fear of fire sales, illiquidity seeking, and credit freezes, *Quarterly Journal of Economics* 76, 557–591.

Diamond, Douglas W., 1984, Financial intermediation and delegated monitoring, *The Review of Economic Studies* 51, 393–414.

Dyck, Alexander, and Lukasz Pomorski, 2016, Investor scale and performance in private equity investments, *Review of Finance* 20, 1081–1106.

Ewens, Michael, Charles Jones, and Matthew Rhodes-Kropf, 2013, The price of diversifiable risk in venture capital and private equity, *Review of Financial Studies* 26, 1853–1889.

Fitzpatrick, Nick, 2011, Insurance companies: investing their way to solvency, http://www.funds-europe.com/insurance-companies-investing-their-way-to-solvency.

Giommetti, Nicola, and Morten Sørensen, 2019, Optimal allocation to private equity, Working paper, Copenhagen Business School.

Glode, Vincent, and Richard Green, 2011, Information spillover and performance persistence for hedge funds, *Journal of Financial Economics* 101, 1–17.

Gompers, Paul, and Josh Lerner, 1996, The use of covenants: An analysis of venture partnership agreements, *Journal of Law and Economics* 39, 463–498.

Griffith, Erin, 2009, Lp defaults: What exactly happens?, https://www.pehub.com/lp-defaults-what-exactly-happens/.

Haddad, Valentin, Erik Loualiche, and Matthew Plosser, 2017, Buyout activity: The impact of aggregate discount rates, *Journal of Finance* 72, 371–414.

Harris, Robert S., Tim Jenkinson, and Steven N. Kaplan, 2014a, Private equity performance: What do we know?, *The Journal of Finance* 69, 1851–1882.

Harris, Robert S., Tim Jenkinson, Steven N. Kaplan, and Rüdiger Stucke, 2014b, Has persistence persisted in private equity? evidence from buyout and venture capital funds, Working paper.

Hochberg, Yael V., Alexander Ljungqvist, and Annette Vissing-Jörgensen, 2014, Informational holdup and performance persistence in venture capital, *The Review of Financial Studies* 27, 102–152.

Holmström, Bengt, and Jean Tirole, 1997, Financial intermediation, loanable funds, and the real sector, *The Quarterly Journal of Economics* 112, 663–691.

Huether, Niklas, David T. Robinson, Soenke Sievers, and Thomas Hartmann-Wendels, 2020, Paying for performance in private equity: Evidence from venture capital partnerships, *Management Science* 66, 1756–1782.

Innes, Robert D., 1990, Limited liability and incentive contracting with ex-ante action choices, *Journal of Economic Theory* 52, 45–67.

Jensen, Michael, 1989, The eclipse of the public corporation, *Harvard Business Review* 67, 61–74.

Kaplan, Steven N., and Antoinette Schoar, 2005, Private equity performance: Returns, persistence, and capital flows, *The Journal of Finance* 60, 1791–1823.

Kaplan, Steven N., and Per Strömberg, 2009, Leveraged buyouts and private equity, *Journal of Economic Perspectives* 23, 121–146.

Kleymenova, Anya, Eli Talmor, and Florin Vasvari, 2012, Liquidity in the secondaries private equity market, Working paper, London Business School.

Laux, Christian, 2001, Limited-liability and incentive contracting with multiple projects, *Rand Journal of Economics* 32, 514–526.

Lerner, Josh, and Antoinette Schoar, 2004, The illiquidity puzzle: theory and evidence from private equity, *Journal of Financial Economics* 72, 3 – 40.

Lerner, Josh, Antoinette Schoar, and Wan Wongsunwai, 2007, Smart institutions, foolish choices: The limited partner performance puzzle, *Journal of Finance* 62, 731–764.

Litvak, K., 2004, Governing by exit: Default penalties and walkaway options in venture capital partnership agreements, *Willamette Law Review* 40, 771–812.

Ljungqvist, Alexander, Matthew Richardson, and Daniel Wolfenzon, 2020, The investment behavior of buyout funds: Theory and evidence, *Financial Management* 49, 3–32.

Longstaff, Francis, 2004, The flight-to-liquidity premium in u.s. treasury bond prices, *Journal of Business* 77.

Lynn, Alex, 2020, Eaton: Large us institutions grappling with capital calls, https://www.privateequityinternational.com/eaton-large-us-institutions-grappling-with-capital-calls/.

Marquez, Robert, Vikram Nanda, and Deniz Yavuz, 2014, Private equity fund returns and performance persistence, *Review of Finance* 19, 1783–1823.

Metrick, Andrew, and Ayako Yasuda, 2010, The economics of private equity funds, *Review of Financial Studies* 23, 2303–2341.

Mitchenall, Toby, 2020, Investor liquidity: Reading the runes, https://www.privateequityinternational.com/investor-liquidity-reading-the-runes/.

Nadauld, Taylor, Berk Sensoy, Keith Vorkink, and Michael Weisbach, 2018, The liquidity cost of private equity investments: Evidence from secondary market transactions, *Journal of Financial Economics* forthcoming.

Phalippou, Ludovic, and Mark M. Westerfield, 2014, Capital commitment and illiquidity risks in private equity, Working paper, University of Oxford Said Business School.

Robinson, David, and Berk Sensoy, 2013, Do private equity managers earn their fees? compensation, ownership, and cash flow performance, *Review of Financial Studies* 26, 2760–2797.

Robinson, David, and Berk Sensoy, 2016, Cyclicality, performance measurement, and cash flow liquidity in private equity, *Journal of Financial Economics* 122, 521–543.

Sahlman, William, 1990, The structure and governance of venture-capital organizations, *Journal of Financial Economics* 27, 473–521.

Silveira, Mark, and David Harris, 2016, When investors default, https://www.mjhudson.com/newsletter/alternative-insight-july-2016/.

Sørensen, Morten, Neng Wang, and Jinqiang Yang, 2014, Valuing private equity, *The Review of Financial Studies* 27, 1977–2021.

Witkowsky, Chris, 2020, Lps feel liquidity heat but no fears of widespread defaults, https://www.buyoutsinsider.com/lps-feel-liquidity-heat-but-no-fears-of-widespread-defaults/.

Zeisberger, Claudia, Michael Prahl, and Bowen White, 2017, *Mastering Private Equity* (Wiley).