Leaving Money on the Table: A Theory of Private Equity Fund Returns *

Robert Marquez  Vikram Nanda  M. Deniz Yavuz
Boston University  Georgia Tech  Purdue University

May 26, 2011

*Contact: Deniz Yavuz (myavuz@purdue.edu). We would like to thank Philip H. Dybvig, Alex Edmans, Paolo Fulghieri, Thomas F. Hellman, Steven N. Kaplan, Laura Lindsey, Umit Ozmel, Merih Sevilir and seminar participants at Arizona State University, NYU, UC Berkeley, UC Davis, the 6th Annual Corporate Finance Conference (2009) at Washington University in St. Louis, and the Entrepreneurial Finance and Innovation Conference (2010).
Leaving Money on the Table: A Theory of Private Equity Fund Returns

Abstract

Evidence indicates that private equity funds, unlike mutual funds, deliver persistent abnormal returns and that top performing funds are often oversubscribed. Why do private equity funds appear to leave money on the table, rather than, say, increasing fund size or fees? We argue that private equity funds are fundamentally different from mutual funds because their success is contingent on matching with high quality entrepreneurial firms, and these firms are looking to match with high ability managers. When entrepreneurs cannot distinguish managers’ ability to add value from managers ability to select good firms, fund manager exerts more effort to select firms. They do this in order to manipulate entrepreneurs’ beliefs about managerial ability to add value, even though firms are not fooled in equilibrium. The model provides several novel time series and cross sectional predictions about performance persistence, fund structure in addition to addressing the questions raised above.

Keywords: venture capital, private equity, performance persistence, signal jamming, fund size, fund fees.
1 Introduction

Anecdotal evidence abounds that some successful private equity funds do not accept all the money that investors are willing to provide.\(^1\) At the same time, these funds appear to generate persistent abnormal returns for their investors (see, e.g., Kaplan and Schoar, 2005, or Phalippou and Gottschalg, 2008),\(^2\) in contrast to the case of mutual funds (e.g., Jensen, 1968).\(^3\) Kaplan and Schoar (2005) conjecture that top-performing funds voluntarily restrict their size which, given diseconomies of scale, can enhance the returns delivered to investors. The performance persistence of successful private equity (PE) funds raises an obvious question: Why don’t these fund managers capture higher rents by increasing fund size or fees?

To address this question, we focus on a fundamental difference between private equity and mutual funds. Unlike mutual funds that invest in public securities, investments by private equity funds are subject to a two-sided matching problem:\(^4\) Private equity funds want to match with high quality entrepreneurial firms; on the other side, entrepreneurs want to match with talented fund managers that are more likely to add value.\(^5\) Hsu (2004) provides evidence that firms are more likely to accept offers from VCs that are more reputable and presumably have a higher ability to add value. Although a fund manager’s ability to add value is not directly observed, inferences can be drawn from returns to past investments. However, it is not always clear whether past returns are a result of a VC’s ability to add value or stem from

\(^1\)Several cases in which VC funds were oversubscribed are noted in the article “Oversubscribed,” European Venture Capital Journal, November 2006. Also see Kaplan and Schoar (2005).

\(^2\)See also Quigley and Woodward (2002), Jones and Rhodes-Kropf (2003), Ljungqvist and Richardson (2003), Cochrane (2005), and Korteweg and Sorensen (2008), which explore the risk, cash flow and performance of private equity funds.

\(^3\)For successful funds, Carhart (1997) shows that “common factors in stock returns and investment expenses almost completely explain persistence in equity mutual funds’ mean and risk-adjusted returns” (p. 57). The worst performing mutual funds do seem to exhibit some performance persistence (see Carhart, 1997), possibly resulting from inattention by investors in these funds.

\(^4\)Matching is important also in financial intermediation (Chemmanur and Fulghieri, 1994; Fernando, Gatchev and Spindt, 2005).

\(^5\)There are a variety of ways in which private equity managers can add value: through providing strategic advice, helping to professionalize firm management, and by attracting better resources, business partners and human capital (Gorman and Sahlman, 1989, Megginson and Weiss, 1991, Hellmann and Puri, 2000, 2002, Baum and Silverman, 2004, and Hochberg, Ljungqvist, and Lu, 2007).
his ability to select and match with high quality firms. Obviously, the ability to add value and deal flow are not independent from each other, which makes it difficult to separate the two effects. Indeed, Sorensen (2007) provides evidence that both effects are important in explaining VC fund success. Our premise is that these information asymmetries about the source of returns provide incentives for fund managers to manipulate entrepreneurs’ beliefs about their ability to add value in order to match with better entrepreneurs in the future.

A private equity fund is raised in several stages and the fund raising process is governed by a series of contractual provisions. Typically, fund managers (general partners) set fund fees and a target fund size before going on a road show to attract capital from institutional investors. During the fund raising process, which may take up to a year, limited partners who enter late are almost always required to pay the same fees as limited partners who entered earlier (Lerner, Harydmon and Leamon, 2007). At the end of this process the fund can be over- or under-subscriberd compared to the initial target size. Investors’ interest during the fund raising process may reveal collective information of their assessment of market conditions and future potential returns. For example, the manager’s inability to raise the minimum fund size may reveal adverse information that the initial investors did not have at the time the fund raising was started (Lerner, Harydmon and Leamon, 2007). In these case, the fund fails to form and investors’ capital commitments are usually cancelled. If the fund is oversubscribed, in many cases the general partner is allowed to increase the fund’s size (Lerner, Harydmon and Leamon, 2007).

We present a model that incorporates these features. In the first stage, the manager sets up a fund by determining what fees to charge. During the fund raising process, all parties learn more about market conditions and in particular about the average quality of available investments. Investors are competitive and are willing to supply funds until their expected returns are zero. After observing investor demand and market conditions, the manager decides on the fund’s size and exerts effort to select firms. Fund returns are jointly determined by the quality of firms in the fund’s portfolio and the manager’s ability to add value. Future entrepreneurs observe returns to fund investors and/or to portfolio firms and
use this information to estimate a manager’s ability to add value. Managers with higher perceived ability attract better entrepreneurs. We first show that when all information about a manager’s effort is available to entrepreneurs, the model collapses to a variant of the model in Berk and Green (2004), with the manager increasing the fund’s size until investors’ expected excess return is zero.

We then move to a more realistic setting in which the fund manager’s effort is not observed. In this case, the manager exerts effort to select higher quality firms in the hope of manipulating the beliefs of entrepreneurs about his ability to add value. Since it is less costly to improve the quality of firms in a smaller portfolio, the manager has a tendency to keep the fund size small. This leads to positive expected returns for investors if the fund’s fees do not fully capture the excess returns generated. While fees are set optimally before the fund raising process anticipating the possibility that the fund’s size may be kept small, extracting all the surplus through fees is not optimal as higher fees lead to a further reduction in the expected fund’s size. They also increase the probability that the fund will not be established at all, since higher fees make it more difficult to satisfy investors’ participation constraint. Since the manager faces a similar tradeoff the next time he sets up a fund, this results in performance persistence over time.

Entrepreneurs are not fooled in equilibrium since they recognize the manager’s incentives and can rationally anticipate his actions. Nevertheless, the manager cannot avoid the attempt to manipulate entrepreneurs’ beliefs since entrepreneurs, in assessing the manager’s ability, assume that the manager is indeed attempting to manipulate their beliefs. As a result, if the manager were to not exert extra effort, the fund would deliver lower returns and these would be regarded by entrepreneurs as being driven by a lower managerial ability to add value. The fund manager would prefer to commit not to exert too much effort, but is unable to do so to the extent that he cannot credibly communicate his true ability, or because he himself does not know it and effort is private.

We extend our model to show that our basic results are robust to other modelling choices.

---

6 This is referred to as a “signal jamming” equilibrium in which an agent tries to affect the principal’s perception of his ability by manipulating the signal. See, e.g., Holmstrom (1999) or Stein (1988).
For instance, our results hold regardless of whether entrepreneurs observe fund returns or returns to individual firms in the portfolio. One also obtains similar results if effort affects fund returns through other channels, such as by mitigating the cost of allocating capital to different investments. Our results are also robust to the inclusion of an option-like performance fee (carried interest or variable fee) in addition to a fixed fee (management fee), or to the possibility that the fund may be disbanded if a minimum size cannot be achieved. In general, similar results could be obtained if there are other unobserved fund characteristics that can affect fund returns and can be manipulated by the manager.

In addition to explaining the performance persistence puzzle, our model provides several unique untested predictions. In general, we expect higher performance persistence when managers’ incentives to manipulate the beliefs of entrepreneurs are higher. This happens when the impact of a manager’s perceived ability to add value on matching is higher, when the impact of selection effort on the quality of the firms in the portfolio is greater, when selection effort is less scalable, and when there is higher uncertainty about a manager’s ability to add value. We also predict a positive correlation of realized excess returns across funds due to the common shock to the pool of investments available. Note that this prediction only arises when managers have incentives to manipulate the beliefs of entrepreneurs and, absent such incentives, we expect no such correlation. We also expect more funds to be established when the common shock is positive, with a subsequent poor performance of new entrants in raising consecutive funds. Managers who experience a larger shock both raise larger funds and provide higher returns to investors.

Testing our predictions requires proxies for the variables discussed above. We speculate that these predictions could be tested using cross-sectional differences among private equity funds in terms of their focus on early versus later investment stage (Gompers and Lerner, 1999), on lead versus non-lead position, on investing in firms versus other funds, and LBO versus VC firms. The matching process we study should be more important for venture capital funds that invest in early stage companies, given that a VC’s added value should be more pronounced in the first two rounds of financing (Chemmanur, Krishnan and Nandy, 2000).
We also expect matching to be more important for lead funds given that they are more likely to be the ones adding value. We expect our arguments to be more applicable for VC funds compared to LBO funds (as shown in Kaplan and Schoar, 2005) given that VC investment in general is less scalable compared to LBO funds (Metrick and Yasuda, 2010). We also expect the matching process and our arguments to be more important for funds that directly invest in portfolio firms rather than other private equity funds.

A contemporaneous paper by Hochberg, Ljungqvist and Vissing-Jorgensen (2010) provides an alternative explanation based on increasing bargaining power of fund investors over time (as in informed bank lending models, see, e.g., Rajan, 1992). Institutional investors initially have no bargaining power; however, they obtain “soft information” concerning the fund manager’s skill after they invest. In the second period, other investors do not observe fund returns and refuse to invest if existing investors do not reinvest. As a result, institutional investors’ collective bargaining power increases in the second period compared to the first period, which results in higher fund returns. Our explanation is very different and based on differences in the investment side rather than the liability side, which makes our cross sectional predictions related to fund focus and type unique. We let the fund managers charge a management fee, which generates a conflict of interest between fund managers and investors regarding the optimal size of the fund and allows us to address why managers do not increase fund size, in addition to fees, to capture higher surplus. Why fund managers voluntarily restrict size is an important element of this puzzle (Kaplan and Schoar, 2005).

In addition, our predictions can explain related existing empirical evidence. For example, Kaplan and Schoar (2005) find that fund size and fund sequence are negatively correlated with fund returns when they control for private equity firm fixed effects (Table VI). This is consistent with our time series prediction that as entrepreneurs learn more about a manager’s ability to add value, fund size increases, fees increase (see, e.g., Gompers and Lerner, 1999) and investors’ expected returns decrease. Existing theories cannot explain this time series evidence simultaneously with performance persistence; theories that assume increasing bargaining power of investors predict that follow-on funds should have higher returns.
than those of the first fund. Kaplan and Schoar (2005) also find that funds that are raised during boom times are less likely to raise follow-on funds, which is consistent with our predictions. Moreover, the relation between institutional/large investors and fund performance has been investigated in some recent papers. Busse, Goyal, and Wahal (2009), for instance, do not find performance persistence in mutual funds that cater to institutional investors and Phalippou (2009) finds that venture capital funds that are expected to be backed by more skilled investors show no performance persistence. These findings suggest that fundamental differences between private equity and mutual funds other than differences in the types of investors may also be important in accounting for performance differences.

Another related paper is Glode and Green (2010), who offer an explanation for performance persistence in hedge funds. In their model, hedge fund managers limit their size and deliver excess returns to investors to induce them to not divulge information about the fund’s strategy. Glode and Green (2010) emphasize that their approach better captures features of the hedge fund environment, such as concerns for confidentiality. This could be less of a concern in private equity if investment strategy is harder to replicate. Several recent papers also analyze the optimal size of venture capital firms. Fulghieri and Sevilir (2009) show that a venture capitalist may limit fund size when it is important to provide entrepreneurial incentives to firms. A small portfolio increases the value-added to each firm by the VC and encourages entrepreneurs to exert higher effort, while a large portfolio allows the VC to reallocate resources in the case of startup failure and to extract greater surplus from entrepreneurs. Inderst, Mueller and Muennich (2007) argue that limiting the size benefits the venture capitalist by decreasing the bargaining position of portfolio firms. We provide an additional reason for why VC funds may limit size, which is to facilitate manipulating beliefs of entrepreneurs through higher selection effort as a way of improving the quality of firms with which the manager can match in the future.
2 Model

At the start of any period $t$, a VC fund manager establishes a new fund, raises cash from investors, and invests in entrepreneurial firms. The payoffs to these investments, distributions to investors, and fund liquidation occur at the end of the period. The manager is then expected to start a new fund in period $t + 1$. The fund raising process involves two steps. In the first step, the manager prepares a private placement memo for investors which specifies the targeted fund size and the fund’s fees $f_t$ as a fraction of total fund size. Given the fund’s stated structure, investors then make commitments to the fund (“committed capital”). However, anecdotal evidence indicates that the fund raising process takes anywhere from 4 months to 1 year. Presumably, during this time period the manager learns more about market conditions and the perception of the investment community about the fund’s potential returns. In the second step, after observing market conditions the manager chooses his effort level and determines the actual fund size (“invested capital”). Fund returns are realized at the end of time $t$.

On the investment side, entrepreneurs need to raise a fixed amount of financing, normalized to $1$, in return for giving a fraction of the company to the investing VC fund. The quantity of shares and the price at which a firm’s shares are sold affect how any value created is shared between the fund and the entrepreneur. We abstract from the details of the bargaining between the entrepreneurs and the fund manager and assume that both the entrepreneurs and the fund emerge with an equity stake in the firms. Hence, both parties receive a strictly positive share of the value created. This, in turn, will lead to matching that is positive assortative: managers with higher perceived ability are more likely to match with good entrepreneurs. Positive assortative matching is a common prediction of various models (Titman and Trueman, 1986, Chemmanur and Fulghieri, 1994, Fernando, Gatchev and Spindt, 2005), and we simply take it as given here. However, it is straightforward to see how positive assortative matching could arise in our setting, and we discuss this in more detail.

7The assumption that the number of periods is infinite simplifies the exposition since it makes the value function stationary. Assuming either a finite number of periods or that there is a probability each period of the manager exiting permanently does not qualitatively affect our results.
We denote a fund manager’s ability or talent by $X$, which is distributed according to $N(\bar{X}, \sigma_X^2)$. The fund’s investment coupled with the fund manager’s ability results in a value added of $X + \varepsilon_t X$ percent in period $t$, where $\varepsilon_t X$ is a manager-specific random shock. The shock $\varepsilon_t$ is i.i.d. over time and is distributed as $N(0, \sigma^2_t)$. The actual realization of $X$ is unknown to the fund manager, fund investors and entrepreneurs. However, market participants update their expectations of the manager’s ability, $E_t[X]$, over time, using the information available. Since managerial ability is associated with a higher value creation and entrepreneurs capture some of this value, they always want to match with the manager that has the highest expected value added, $E_t[X]$, among available funds. We assume that the per unit cost of adding value is $S(Q_t)$, where $Q_t$ is the amount of invested funds. This cost is independent of ability and is increasing and convex in $Q_t$ (this is as in Berk and Green, 2004), and is borne by the fund, so that it does not affect returns to entrepreneurs. This assumption implies decreasing returns to scale of adding value to the private equity fund, as documented by Lopez-de-Silanes, Phalippou and Gottschalg (2009). The gross percentage value added for a dollar of investment is proportional to $W_t = X + \varepsilon_t X - S(Q_t)$.

The return to the fund depends not only on the fund manager’s ability to add value but also on the quality of firms in the portfolio. In any period $t$, there is a continuum of firms/entrepreneurs\(^8\) of varying intrinsic quality. A higher quality firm is defined as one for which a fund’s investment, coupled with the fund manager’s ability, leads to higher value creation. The expected (average) quality of firms available for investment, $\bar{P}$, is common knowledge at the beginning of the period. However, at the end of the fund raising process all parties learn more about the average quality of investments available. We model this by assuming that the shock to the average quality of firms available for investment is equal to $\varepsilon_t^P$, where the shock $\varepsilon_t^P$ is i.i.d. over time and is distributed as $N(0, \sigma^2_t)$. The random shock is unknown at the time when the fund manager sets up the fund but becomes known during fund raising. This information may arise due to changes in macroeconomic conditions (fund

\(^8\)We will use the terms “firm” and “entrepreneur” interchangeably to refer to the party receiving an investment from a VC fund.
raising often takes up to a year) or partners may learn from each other by observing each others’ interest in the fund. For example, we know that if there is not sufficient demand for the fund, i.e., if the manager cannot raise more than the minimum target fund size, initial investors who signed up may ask for their money back because of potential negative information others could have (Lerner, Harydmon and Leamon, 2007). This uncertainty in the fund raising process may result in over- or under-subscription compared to the initial target fund size. Once $\varepsilon_t^P$ becomes common knowledge at the end of the fund raising process, investors determine how much capital they are willing to supply and the manager sets the fund’s size.

The final (realized) quality of firms in the fund’s portfolio depends on the average quality of firms available for investment, the shock to the average quality, the perceived ability of the manager to add value, and the manager’s effort to select firms. We define $P_t = z(P,e_t,E_t[X]) + \varepsilon_t^P$ as the average quality of firms in the manager’s portfolio, where $e_t$ is the managerial effort to select firms. In particular, the quality of firms in the portfolio is an increasing and concave function of effort, and is an increasing function of the expected ability to add value by the fund manager. This implies $\frac{\partial P_t}{\partial e_t} > 0$, $\frac{\partial^2 P_t}{\partial e_t^2} < 0$, and $\frac{\partial P_t}{\partial E_t[X]} > 0$. The cost of effort $C$ is increasing and convex in effort and is increasing in the size of the manager’s portfolio: $\frac{\partial C}{\partial e_t} > 0$, $\frac{\partial^2 C}{\partial e_t^2} > 0$, $\frac{\partial C}{\partial Q_t} > 0$, and $\frac{\partial C^2}{\partial Q_t \partial e_t} > 0$ for $e_t > 0$ and zero otherwise.\footnote{Note that all these conditions are satisfied for simple cost functions such as $e_t^2 Q_t$.} A cost of effort that increases with respect to fund size captures limited resources that partners can allocate on screening firms.

The gross percent return to the fund at the end of the investment period is equal to $P_t W_t$. The net return to fund investors after fees is $\alpha_t = P_t W_t - f_t$. Investors are competitive and risk-neutral and, hence, are willing to provide capital as long as their expected net return, $E_t[\alpha_t]$, is nonnegative.

The manager chooses his actions to maximize his expected payoff subject to the participation constraint of the investors and the impact of his actions on entrepreneurs’ beliefs about his ability $X$. We can write the total expected future payoff for the fund manager as
of time $t$ as

$$ V_t = E_t \left[ \sum_{i=t}^{\infty} \delta^i (Q_i f_i - C_i) \right], $$

where $\delta < 1$ is the discount factor, which we ignore in the rest of the paper for brevity.

It is useful to summarize the differences and commonalities in the information sets of investors, entrepreneurs and the manager. All fund characteristics including fund size, fees and functional forms for cost of effort, cost of generating value and returns are common knowledge. The realized fund return and the random shock to the average quality of firms are also common knowledge at the end of the investment period. We use fund’s history $h_t$ at time $t$, to summarize common knowledge to fund investors and entrepreneurs, which always includes all realized returns $\alpha_i$ an fund characteristics for $i < t$. On the other hand, no one knows the manager’s ability to add value $X$ and the per period shock $\varepsilon_t^X$. The only difference between the fund manager and outsiders (entrepreneurs and limited partners) is that the manager always knows his effort. As a result, entrepreneurs who observe $P_t W_t$ cannot disentangle whether higher returns comes from higher effort (therefore higher $P_t$) or higher ability to add value (and therefore higher $W_t$).

### 3 The Symmetric Information Case

As a starting point, we analyze the case where there are no information asymmetries between the fund manager and entrepreneurs. All parties observe managerial effort and the initial quality of firms in the fund’s portfolio, which allows them to easily back out the manager’s ability to add value. While not very realistic, assuming perfect and symmetric information helps us to emphasize the importance of private information later. The symmetric information case is similar to that studied in Berk and Green (2004), who show that there is no performance persistence in mutual fund returns when managers receive a fee that is a fraction of assets under management. This occurs because positive returns attract fund flows, which in turn eliminate a manager’s ability to provide excess returns. As we show below, a similar result obtains in our setting.
We solve the model using backwards induction. We first solve for entrepreneurs’ update of managerial talent at the end of the period, after returns from investments are realized. Specifically, in each period entrepreneurs observe the manager’s choices of fund size $Q_i$, fees $f_i$, managerial effort $e_i$, and time varying quality of firms in the market $\epsilon_i^P$, for $i \leq t$. Entrepreneurs can then use the information available to them to estimate a manager’s ability to add value, $X$. Bayesian updating after observing the total return $P_t W_t$ gives us the conditional expectation of $X$ at time $t+1$ as:

$$E_{t+1} [X|h_t] = w_t E_t [X|h_{t-1}] + (1 - w_t) \left( \frac{P_t W_t}{P_t} + S(Q_t) \right).$$

(1)

Since $W_t = X + \epsilon_t^X - S(Q_t)$, (1) simplifies to

$$E_{t+1} [X|h_t] = w_t E_t [X|h_{t-1}] + (1 - w_t) \left( X + \epsilon_t^X \right).$$

The weights $w_t$ reflect the importance of the most recent realization of fund returns relative to the past history, $h_{t-1}$, in updating entrepreneurs’ expectations concerning $X$. The weight assigned to new information, $1 - w_t$, decreases over time through simple Bayesian updating.\(^\text{10}\)

After investor demand to the fund has been observed and $\epsilon_i^P$ is revealed, the manager then decides on his effort level and fund size. Given that managerial actions have no effect on entrepreneurs’ beliefs about the manager’s ability to add value in future periods, the manager simply maximizes his payoff for the current period. The maximization problem becomes

$$\max_{Q_t, e_t} Q_t f_t - C_t$$

\(^{10}\)Given that the underlying distributions are normal, the weight placed on the new information in Bayesian updating is determined by the precision of the new information relative to the precision of the prior. Specifically, if the conditional distribution of $X$ at $t - 1$ is denoted by $N(E(X|h_{t-1}), \sigma_X^2 | t-1)$, we have $1 - w_t = 1 - \frac{1/\sigma_t^2}{1/\sigma_{t-1}^2 + 1/\sigma_X^2 | t-1}$. It is apparent that as the precision of the conditional distribution of $X$ increases through time, the weight placed on the new information will correspondingly decrease.
subject to the participation constraint of fund investors,

\[ E_t[\alpha_t | e_t^P] = P_t E_t[W_t] - f_t \geq 0. \]

**Proposition 1** When there is no information asymmetry between the fund manager and entrepreneurs, for any given \( f_t > 0 \) the manager chooses a fund size \( Q_t \) and effort \( e_t \) such that fund investors’ expected excess return, \( E_t[\alpha_t | e_t^P] \), is zero. There is also a value of uncertainty \( e_t^P \) such that no fund is raised for \( e_t^P < e_t^P \). In this case, all parties’ expected return is zero.

This result is essentially the same as in Berk and Green (2004). When there is no information asymmetry between the fund manager and entrepreneurs, the manager captures all the surplus by taking all the funds that investors are willing to provide given manager’s (optimal) effort level. Even for very high realizations of \( e_t^P \), the manager can extract all the surplus by increasing fund size and decreasing effort. Effort level is bounded below by zero; however, there is no upper bound on fund size (as long as investors are willing to supply capital) and decreasing returns to scale ensures that the participation constraint of investors will eventually bind as fund size increases. The manager exerts some effort to screen firms, which helps to ease the participation constraint of investors and increases the size of the fund even further.

We now consider the manager’s problem in determining the optimal fee in the first stage given that he knows he will capture all the surplus in the second stage. Again, the manager maximizes his surplus for the current period given that his choice of fixed fee has no effect on entrepreneurs’ expectations about his ability. The maximization problem is

\[ V_t(E_t[X|h_{t-1}]) = \max_{f_t} E [Q_t f_t - C_t], \]

subject to fund size \( Q_t \) and effort \( e_t \) being chosen optimally (i.e., as the solutions to (2) in Proposition 1).
Proposition 2 The optimal fee $f_t^* < \infty$ balances larger fees per dollar under management with lower fund size, higher cost of effort and lower probability of establishing the fund.

The manager considers all implications of fees on his payoff. For a given fund size, his payoff increases as fees increase, and the equilibrium level of effort increases as well. However, the optimal fund size is decreasing in the charged fees because the participation constraint of investors binds in the next stage for higher fees (see Proposition 1). To ease the participation constraint, the manager commits to a higher level of effort, which is costly and is observed by investors given we posit that all information is symmetric. Finally, a higher level of fees increases the probability that investors’ participation constraint cannot be satisfied for any combination of fund size and effort, so that the fund cannot be established.

4 The Role of Asymmetric Information

We next analyze the more realistic case where the VC’s effort to select firms is not observed. In this case the manager may have an incentive to increase his effort in order to improve entrepreneurs’ perceptions concerning his ability to add value.

At the end of the investment period total return in time $t$, $P_t W_t$, is observed. However, entrepreneurs do not know the manager’s effort level $e_t$, which implies that they also do not know the realized quality of the portfolio, $P_t$. This prevents them from perfectly backing out the value $X + \varepsilon_t^X$. Entrepreneurs need to conjecture the actual level of managerial effort in order to assess the manager’s ability. We denote such conjecture as $e_t^C$. As before, entrepreneurs’ form their expectations about the manager’s ability based on the history of realized returns for firms. Bayesian updating after observing the value of the total return $P_t W_t$ gives us the conditional expectation of $X$ at time $t+1$:

$$E_{t+1}[X|h_t] = w_t E_t[X|h_{t-1}] + (1 - w_t) \left( \frac{P_t W_t}{P_t(h_t, e_t^C)} + S(Q_t) \right),$$

where we use $P_t(h_t, e_t^C)$ to denote the market’s conjectured average quality of firms in the
fund’s portfolio, given the conjectured effort level $e_t^C$. This expression can be written as

$$E_{t+1}[X|h_t] = w_tE_t[X|h_{t-1}] + (1 - w_t) \left( \frac{P_t(h_t, e_t)(X + \epsilon_t^X - S(Q_t))}{P_t(h_t, e_t)} + S(Q_t) \right). \quad (4)$$

It is obvious from (4) that $\frac{\partial}{\partial e_t} E_{t+1}[X|h_t] > 0$ because $E[X + \epsilon_t^X - S(Q_t)]$ must be positive when the manager charges a positive fixed fee. Otherwise, investors would earn negative returns in expectation and would choose not to participate. Therefore, at the time when the manager decides on his effort level, entrepreneurs’ perception of the manager’s talent is increasing in the actual effort given the conjectured level of effort. This adds an important dimension to the manager’s decision over the fund’s fees, fund size, and his effort level. The manager now has to consider the effects of his decisions on his future payoff through their effect on entrepreneurs’ beliefs about his ability to add value. In other words, the manager’s optimization problem becomes dynamic. For a given fee $f_t$, we can write the Bellman equation of the fund manager as:

$$V_t(E_t[X|h_{t-1}]) = \max_{Q_t, e_t} \left[ Q_t f_t - C_t + V_{t+1}(E_{t+1}[X|h_t]) \right], \quad (5)$$

with the additional constraint that the participation constraint of fund investors must be satisfied:

$$E_t[\alpha_t \epsilon_t^P] = P_t E[W_t] - f_t \geq 0. \quad (6)$$

**Proposition 3** There is an upper bound on fund size, $Q_t^{\max}$, such that the optimal fund size is $Q_t^* \leq Q_t^{\max}$.

Analysis of the first order conditions with respect to fund size and effort establishes an important observation, which is that there is an upper bound on fund size. This is introduced by the manager’s incentive to manipulate entrepreneurs’ beliefs about his ability to add value. In order to achieve this goal the manager always wants to exert some effort, $e_t^{\min}$, regardless of whether the investors’ participation constraint binds or not. The cost of improving the quality of firms in a larger portfolio is greater, which introduces an upper
bound on the quantity that the manager is willing to invest. Note that when there is no information asymmetry, as in Section 3, there is no such upper bound on fund size except the one imposed by investors’ participation constraint.

**Proposition 4** For any given \( f_t \), there is a realization of uncertainty \( \varepsilon_t^{P,\text{low}} \) such that fund investors’ expected excess return \( E_t[\alpha_t|\varepsilon_t^P] \) is strictly positive for \( \varepsilon_t^P > \varepsilon_t^{P,\text{low}} \) and zero otherwise. There is also a value of uncertainty \( \varepsilon_t^P \) such that no fund is raised for \( \varepsilon_t^P < \varepsilon_t^{P,\text{low}} \). In this case, all parties’ expected return is zero.

When the manager has an incentive to manipulate entrepreneurs’ beliefs, he optimally chooses to not capture all the surplus in certain states. Consider the case where \( \varepsilon_t^P \) is positive and large, so that at the target fund size investors’ participation constraint would not bind. In this case the manager could increase the fund’s size (and/or decrease his effort) to capture all the surplus. However, this may not be optimal if the manager has to increase the fund’s size beyond \( Q_t^{\text{max}} \). For fund sizes larger than \( Q_t^{\text{max}} \), manipulating the beliefs of entrepreneurs will be simply too costly. Therefore, in some states of the world (i.e., for some realizations of \( \varepsilon_t^P \)) the manager provides positive expected returns to investors.

The proposition proves that for any given fee \( f_t \) there is some value \( \varepsilon_t^P \) such that the manager provides positive expected returns to fund investors. However, this does not establish that in equilibrium the expected return to investors will be positive since in principle the manager could choose a fee in the first stage such that probability of leaving a positive expected return to fund investors is negligible. The manager’s goal is to maximize his own payoff rather than simply capturing all the surplus. These two goals need not be the same since a larger fee that capture more of the surplus may also decrease the total surplus available.

To formally analyze this trade-off, we solve the manager’s optimization problem in the first stage. As above, the problem is inherently dynamic, although the manager’s choice of the optimal fee does not directly affect entrepreneurs’ beliefs about his ability given that fees
are observed. The fund manager’s objective function is

\[ V_t(E_t [X|h_{t-1}]) = \max_{f_t} \int_{\{\epsilon_{it}^P : P_t E[W_t] - f_t \geq 0\}} (Q_t f_t - C_t) dN (\epsilon_{it}^P), \]  

(7)

where \( N(.) \) is the cdf of \( \epsilon_{it}^P \), which is normally distributed. We define the solution to (7) under asymmetric information as \( f_t^A \).

**Proposition 5** The optimal fee \( f_t^A \) < \( \infty \) balances a larger fee against a lower fund size and a lower probability of establishing the fund. At the optimal fee \( f_t^A \), fund investors’ expected return, \( E_t [\alpha_t] \), is strictly positive. This results in performance persistence over time.

Analysis of the first order condition indicates that the optimal fee is determined by considering a number of trade-offs. A larger initial fee results in a lower fund size when fund investors’ participation constraint binds in the second stage. On the other hand, a larger fee decreases the region where the fund manager leaves a positive excess returns to investors. In some sense, as the manager increases fees he increases the probability that he will capture all the surplus form the fund investors; at the same time, however, this reduces the expected fund size on which he earns this fee. There is also a region of uncertainty about the shock to the average quality of firms, \( \epsilon_{it}^P \), where a fund may not be established, and the size of this region increases with the fee \( f_t \). As a result, the optimal fee is determined by balancing these effects and it is optimal to set it in a way that does not capture all the surplus for all realizations of \( \epsilon_{it}^P \). Fund investors’ expected return is therefore positive in equilibrium.

We showed that the fund manager provides positive expected return to fund investors when there is information asymmetry about his ability to add value, which gives the manager an incentive to attempt to manipulate the beliefs of investors. This also translates into performance persistence over time given that the manager is faced with a similar problem every time he establishes a new fund. Performance persistence here is defined as the provision of positive expected excess returns to investors in every fund that is established by the same manager. This expected performance persistence translates into the empirically observed performance persistence across a large number of observations.
4.1 Comparative Statics and Empirical Predictions

We next analyze comparative statics of our main result, both in the time series and in the cross section. We focus on the asset side of private equity to explain differences in performance persistence between private equity and mutual funds. In that regard, to the best of our knowledge all of our cross sectional predictions are unique to our framework. We first summarize the direct predictions of the model by analyzing the effect of different parameters on performance persistence. Next we speculate on how these predictions could be tested and how these tests could be used to differentiate our theory from others.

We start by rephrasing a result from the proof of Proposition 4 that helps illustrate how changing some of the model parameters affects investors’ expected returns. As we have shown, investors’ expected return $E_t[\alpha_t]$ at the time a fund is established is positive. After the realization of $\varepsilon_t^P$, the expected return will be positive when $\varepsilon_t^P > \varepsilon_t^{P,\text{low}} = \frac{f_t}{E[W(Q_{\text{max}}^t)]} - z(e_{\text{min}})$, and zero otherwise. Intuitively, as this cutoff value $\varepsilon_t^{P,\text{low}}$ decreases, the probability that investors’ expected return is positive at the time a fund is established increases. This relation will be useful to determine the effect on performance persistence of changes in various parameters of the model.

Positive assortative matching in private equity (Sorensen, 2007) and its impact on the incentives of managers to manipulate the beliefs of investors is an essential component of our model. We predict that as positive assortative matching becomes more important, we should observe a higher probability of positive expected excess returns and higher performance persistence. This happens because as matching becomes more important managers have higher incentive to manipulate beliefs of investors. Managers exert higher effort and decrease the expected fund size and fees, which results in higher returns to fund investors.

The matching process becomes more important under the following conditions. If there is higher uncertainty about managers ability to add value this increases managers incentive to exert higher effort to manipulate beliefs of investors. However, over time, investors learn more about a manager’s ability to add value. Therefore, the weight they assign to recent returns in updating their beliefs about the managers ability, $1 - w_t$, will be lower due to Bayesian
updating. This decreases the manager’s incentive to manipulate the beliefs of investors by affecting the impact of effort on the expected ability of the manager. As a consequence, the manager will exert lower effort and will choose a larger maximum fund size, which in turn reduces the probability of positive expected returns for investors. In addition, if $\frac{\partial P_t}{\partial E_t[X]}$ is higher, managers have higher incentive to try to manipulate the beliefs of entrepreneurs because that has higher effect on the quality of the portfolio in the future. This increases equilibrium level of effort and decreases both expected fund size and maximum fund size. This in turn results in higher expected return for fund investors. The same mechanism applies to all other variables that affect the managers’ effort in the same direction. Similarly, if the impact of effort in selecting firms is higher, managers exert more effort to manipulate the beliefs of investors. On the other hand, if the cost of effort is higher then a fund manager will exert lower effort. If the manager’s cost of effort increases quickly with the size of the fund, the manager is less likely to increase the fund’s size to capture more of the surplus. In other words, if selection effort is not easily scaled we expect the manager to choose a lower maximum fund size and investors’ probability of receiving positive expected return increases. A manager’s manipulation has higher impact on his payoff if his expected ability to add value is higher. The quality of the firms in the portfolio is multiplied by the manager’s ability to arrive at the total value generated. This complementarity of selection effort and value adding ability encourages a manager with a higher ability to spend more effort in selecting firms. These are summarized in the following corollary.

**Corollary 6** As positive assortative matching becomes more important we should observe a higher probability of positive expected excess returns and higher performance persistence. The positive assortative matching becomes more important:

1) If uncertainty about a manager’s ability to add value increases, i.e., $(1 - w_t)$ increases. Over time, through bayesian updating, $(1 - w_t)$ decreases, and therefore a manager’s incentive to manipulate investors’ beliefs decreases.

2) If the manager’s perceived ability to add value has higher impact on matching, i.e., if $\frac{\partial P_t}{\partial E_t[X]}$ is larger.
3) If selection effort has a greater marginal impact on the quality of firms in the fund’s portfolio, i.e., if $\frac{\partial P_t}{\partial e_t}$ is larger, or if the marginal cost of effort is lower, i.e., if $\frac{\partial C_t}{\partial e_t}$ is lower.

4) If the cost of effort is less scalable, i.e., if $\frac{\partial C_t}{\partial Q_t}$ is larger.

5) If the manager has a higher ability to add value, i.e., a higher $X$.

The realization of uncertainty about the average quality of investments available, $\varepsilon_t^P$, may differ across managers or groups of managers that focus on different investment strategies. This allows us to come up with the following predictions. Although the threshold value $\varepsilon_t^{P,low}$ that results in expected positive returns to fund investors varies across managers, a higher realization of $\varepsilon_t^P$ increases the probability of positive expected return across all funds affected by this shock. Therefore, we predict that fund investors’ expected returns will be positively correlated across funds that are established during similar time periods and with similar investment strategies. This may seem like an obvious prediction. However, we should emphasize that such correlation is not a direct result of the common shock. Indeed, when there is no information asymmetry managers respond to a positive shock by increasing the size of their funds to capture the additional expected surplus, therefore eliminating any correlation of fund investors’ expected return across funds. Assume that we have a cross section of managers with different fund characteristics and ability to add value. With a positive common shock to pool of investments available, more funds can be established because even lower ability managers can now raise funds. This implies that a larger number of funds will be established in that time period. However, the probability that existing managers can raise consecutive funds decreases simply because managers who were able to marginally set up funds with the help of positive common shock are not expected to be able to raise funds next period. Assume that the realization of uncertainty about average quality of investments available, $\varepsilon_t^P$, differs across managers that focus on different investment strategies. Managers who experience a positive shock on average will raise larger funds. These managers are also likely to provide higher returns to investors given that if $\varepsilon_t^P > \varepsilon_t^{P,low}$ fund investors’ expected return is positive. We summarize these in the following proposition.
Corollary 7 1) Positive expected excess fund returns are correlated among funds that are subject to the same shock $\varepsilon_t^P$.

2) When the realization of $\varepsilon_t^P$ is higher, a larger number of funds are established. However, the probability of raising consecutive funds decreases.

3) Managers with higher realization of $\varepsilon_t^P$ raise larger funds and provide higher expected returns to fund investors.

The predictions of our model are consistent with the existing empirical evidence. Kaplan and Schoar (2005) provide evidence regarding the relationship between fund performance, fund sequence and size. In the time series, when they control for private equity firm fixed effects, they find that fund size and fund sequence are negatively correlated with fund returns (Table VI). This is consistent with our time series prediction that as entrepreneurs learn more about a manager’s ability to add value, fund size increases and investors’ expected returns decrease (Corollary 6-1). To the best of our knowledge, existing theories offer no explanation for this finding. For example, theories that assume increasing bargaining power of investors imply that the returns to second fund should be higher than those of the first fund, so that in the time series fund sequence would be positively correlated with fund returns.

In the cross section, we predict that managers that experience a higher $\varepsilon_t^P$ raise larger funds and provide higher expected returns to fund investors (Corollary 7-3). This is consistent with Kaplan and Schoar’s (2005) pooled regression findings, which show that fund size and sequence are positively correlated with fund returns when fixed effects for private equity firms are not included. This finding could also be explained by new entry by a large number of managers who have no ability to add value, together with a few managers that have ability to add value. Given that on average there is some ability, these new managers could raise small funds. The majority of these funds would do very poorly and disappear after investors update their beliefs about the manager’s ability. Managers with ability would survive and would go on to raise larger funds. Assuming that there is new entry every period, a pooled regression would compare managers with no ability to add value who raise small funds with managers who have ability and have larger funds. Indeed, Kaplan and Schoar
(2005) find that, when many new funds enter, overall performance of the industry decreases largely because of the poor performance of new entrants. They also find that funds that are raised during boom times are less likely to raise follow-on funds. This is consistent with our prediction that positive shocks of $\epsilon_t^P$ results more funds to be established; however, the probability of raising consecutive funds decreases (Corollary 7-2).

In addition, our predictions regarding the cross section of funds are unique to our model and could be used to test our theory. For example, there are cross-sectional differences for funds in terms of focus on early versus later investment stage (Gompers and Lerner, 1999), lead versus non-lead position, and investing in firms versus other funds. Private Equity funds also differ in their types, such as venture capital or private equity funds. We next speculate on how our predictions could be tested using these cross sectional differences as proxies for the model parameters discussed above. Empirical evidence finds that VC’s value added is more pronounced in the first two rounds of financing (Chemmanur, Krishnan and Nandy, 2009). This is consistent with the arguments that VCs add value by providing strategic advice, helping to professionalize firm management, and by attracting better resources, business partners and human capital (Gorman and Sahlman, 1989, Megginson and Weiss, 1991, Hellmann and Puri, 2000, 2002, Baum and Silverman, 2004, and Hochberg, Ljungqvist, and Lu, 2007). These services are presumably more valuable at earlier stages. Therefore, the matching process should be more important for VC funds that specialize in earlier rounds of investing. From Corollary 6-2&5 we predict higher performance persistence among VC funds that specialize in investing in earlier rounds compared to funds that specialize in investing in later rounds. Similarly, one could also argue that lead VCs may spend more time with their portfolio firms and potentially add more value compared to non-lead VCs. Therefore, matching with a good lead VC should be more important than matching with other VCs that invest as a part of syndicate but do not directly add value to the firm. Another way of testing the importance of matching on performance persistence could be to look at differences between funds of funds and funds that directly invest in firms. Presumably, the matching process (and value added) should be more important for funds that directly match
with entrepreneurs rather than funds that invest in other funds.

There are also important differences across private equity fund types. For example, buyout firms scale their size much faster in reaction to past positive returns (Metrick and Yasuda, 2010) and their performance persistence appears to be lower (Kaplan and Schoar, 2005, table VII) compared to venture capital funds. Evidence here is consistent with our predictions. Metrick and Yasuda (2010) and others have argued that LBO funds have higher scalability. Higher scalability of selection effort results in larger fund size and lower performance persistence in our model. Moreover in LBO deals, matching with a good manager could be less important for entrepreneurs especially if they sell the entire company for cash. An exception may be LBO funds that keep the management of the target firm in place: although management turnover increases after LBOs, some managers stay with the firm (Muscarella and Vetsuypens, 1990, and Cornelli and Karakas, 2010) and management often has substantial ownership after the LBO (Kaplan, 1989, Kaplan and Stromberg, 2009). Therefore, management may be more supportive of a buy-out offer if the buyout fund has a higher ability to add value.

Finally, we speculate that a manager’s incentive to manipulate entrepreneurs’ beliefs may depend on other fund characteristics that could affect the matching process such as geographic separation (Chen, Gompers, Kovner and Lerner, 2009), specialization (Fulghieri and Sevilir, 2009; Gompers, Kovner, Lerner, Scharfstein, 2008), or fund availability over time. Testing whether performance persistence varies within these dimensions could help us learn more about the importance of the matching process in performance persistence.

5 Model Extensions

Our main intuition is that if a fund manager’s ability to add value is important for matching with high quality entrepreneurs, then managers have incentives to attempt to manipulate the beliefs of entrepreneurs and this may result in smaller fund size and performance persistence. We provided a simple framework that establishes these results. In this section we extend
our model in various directions.

5.1 Observing Returns to Portfolio Firms

In our model, potential future entrepreneurs observe fund returns and try to deduce a manager’s ability to add value. Instead, entrepreneurs could also try to estimate a manager’s ability to add value by observing the value of individual portfolio firms that have gone through an IPO. Sometimes estimating managerial ability by looking at individual firms could be easier if some of the fund’s characteristics are not observed. We note, however, that all of our intuition and proofs go through if investors use returns to portfolio firms rather than the returns to fund investors to draw inferences on managers’ abilities to add value.

To see this, define $P_{it}$ as the intrinsic quality of an individual firm $i$ in the fund’s portfolio that has undergone an IPO.\footnote{The average quality of firms in a fund’s portfolio would then be defined as $P_t = \frac{\sum_i P_{it} W_i}{Q_t}$, which would be equivalent to the measure of portfolio quality we have used so far.} Observing the value of the firm at the IPO allows outside entrepreneurs to directly calculate the returns to entrepreneur $i$. Returns to individual entrepreneurs will be an increasing function of value generated $P_{it} W_i$ and of how this value is shared. But once again, it will not be clear to outsiders whether high observed returns are due to high quality of the firm (due to higher selection effort) or due to the manager’s higher ability to add value. Therefore, a manager has incentive to manipulate the beliefs of investors to maximize his probability of matching with higher quality firms regardless of whether outsiders observe returns to individual firms or fund’s portfolio.

5.2 Managerial Effort that Improves Fund Returns

While the results above are presented in the context of unobserved effort on the part of the manager to select firms, theoretically any unobserved choice variable the manager controls and which affects gross returns can provide similar incentives to the manager to manipulate the beliefs of entrepreneurs. These could be decisions related to unobserved characteristics
of the portfolio firms, fund characteristics or managerial effort. To illustrate this point, we introduce unobserved managerial effort that improves the value of the investments for the fund, for instance by mitigating the cost of allocating funds, without directly affecting the quality of firms in the portfolio. For simplicity, here we ignore the managerial effort to select firms. The gross return delivered by a fund manager in any period \( t \) is now given by

\[ W_t = X + \varepsilon_t^X - (S(Q_t) - o_t), \]

where \( o_t \) denotes the manager’s effort. Effort is privately costly to the manager, with the cost equal to \( K(o_t, Q_t) \), which is increasing and convex in effort, and is increasing with the size of the fund:

\[ \frac{\partial K}{\partial o_t} > 0, \frac{\partial^2 K}{\partial^2 o_t} > 0, \frac{\partial K}{\partial Q_t} > 0, \frac{\partial K^2}{\partial Q_t \partial o_t} > 0 \]

for \( o_t > 0 \) and zero otherwise. Therefore, the average quality of the fund’s portfolio, \( P_t \), is determined only by the perceived ability of the manager. The rest of the model is the same as before.

Since effort \( o_t \) is not observed, entrepreneurs will need to conjecture the actual level of managerial effort in order to assess the manager’s ability. We denote such conjecture as \( o^C_t \). As before, entrepreneurs form their expectations about the manager’s ability based on the history of realized net returns for investors as well as the current returns:

\[ E_{t+1}[X|h_t] = w_tE_t[X|h_{t-1}] + (1 - w_t)(X + \varepsilon^X_t + o_t - o^C_t). \]  

(8)

Given entrepreneurs’ fixed beliefs about the manager’s effort choice, \( o^C_t \), the manager’s actual decision clearly affects the perceived talent next period since

\[ \frac{\partial}{\partial o_t} E_{t+1}[X|h_t] = 1 - w_t > 0. \]

We can write the Bellman equation as

\[ V_t(E_t[X|h_{t-1}]) = \max_{Q_t, o_t} [Q_t f_t - K(Q_t, o_t) + V_{t+1}(E_{t+1}[X|h_t])], \]

subject to the usual participation constraint, \( P_tE_t[W_t] - f_t \geq 0 \), well as (8).

As argued above, asymmetric information about managerial effort implies that \( \frac{\partial E_{t+1}[X]}{\partial o_t} > 0 \) and a greater perceived managerial talent translates into a higher expected continuation
value. This effect is not present when information is symmetric, and it increases the incentive for the manager to exert higher effort because he considers not only the contribution to his payoff in the current period but also the effect of his effort on his future payoff through entrepreneurs’ expectations of his ability. Entrepreneurs will, of course, recognize the manager’s incentive to mislead them. In equilibrium, entrepreneurs’ conjectures must be correct and they will not be fooled by the manager’s effort choice. Nevertheless, the manager attempts to manipulate entrepreneurs’ beliefs by increasing his effort because otherwise he would be perceived as having low ability. In addition, the manager raises a smaller fund than what is optimal when effort is observed, despite the fact that his choice of fund size is observed by the entrepreneurs. This occurs because managing a larger fund makes it more costly to provide higher returns to investors through higher effort. Therefore, our results continue to hold when effort affects fund returns through mitigating costs rather than through improving the quality of firms.

**Proposition 8** For any $f_t$, there is a value of uncertainty $\varepsilon_t^{P,low}$ such that fund investors’ expected excess return $E_t[\alpha_t | \varepsilon_t^P] > 0$ for $\varepsilon_t^P > \varepsilon_t^{P,low}$ and zero otherwise. Therefore, at the optimal fee fund investors’ expected return, $E_t[\alpha_t]$, is strictly positive.

We assumed that the fund characteristics such as fund fees and size are observed. Any uncertainty about these characteristics that are under a manager’s discretion may also present an opportunity to manipulate the beliefs of investors. It is straightforward to extend the model to incorporate such uncertainty. One such unobservable could be fund fees to the extent that outside entrepreneurs observe returns to the fund but not direct returns to all portfolio firms. Although most funds announce a 1.5-3 percent management (i.e., “fixed”) fee and a 20-25 percent carry interest (i.e., “variable” fee) (Gompers and Lerner, 1999; Phalippou and Gottschalg, 2008; Litvak, 2009), in many instances it appears that there is some uncertainty regarding the exact fees that are paid and the discretion the manager has in determining the level of these fees. For instance, there may be hidden fees such as distributions that generate an interest free loan to VCs from limited partners or kickbacks.
that rewards investors. Litvak (2009) argues that the timing of distributions might be as important a part of managerial compensation as the management fee or carry interest. Some fees are not visible to outsiders such as the fees charged to portfolio companies by buyout funds (Phalippou, 2007; Metrick and Yasuda, 2010). There are also a number of extra fees or costs that can be imposed, leading some observers, such as Phalippou (2007), to conclude that fee contracts are opaque. In this case, a manager would charge a lower fee and provide higher returns for fund investors, hoping to manipulate beliefs about his ability to add value.

5.3 Restrictions on Minimum Fund Size

Most private equity partnership agreements require the fund to have a minimum and a maximum size (Lerner, Harydmon and Leamon, 2007). Generally, if the fund is oversubscribed it is allowed to exceed the maximum size with the permission of the limited partners. Obviously, when the fund is oversubscribed the fund manager has a good bargaining position. However, if the fund manager is not able to attract enough capital, initial investors may prefer that the fund be disbanded. This could be because the fund manager’s inability to raise enough funds may signal adverse information that other potential investors have (Lerner, Harydmon and Leamon, 2007) or simply because a minimum fund size is essential to make desired investments. In this section we discuss the implications of such an exogenously imposed lower bound on fund size.

We introduce this lower bound on fund size $Q_{\text{min}}$ by assuming that if at the end of the fund raising process the manager fails to raise this amount of money, the fund is disbanded. If disbanded, the fund manager’s payoff is zero.

**Proposition 9** A lower bound on fund size has two effects on investors’ expected return. First, in some states of the world a lower bound forces the manager to accept a larger fund size than he otherwise would. This negatively affects investors’ expected excess returns. Second, the manager sets a lower fee recognizing the possibility that otherwise the minimum of $Q_{\text{min}}$ may not be raised. This positively affects investors’ expected excess return.
The (endogenous) upper bound on fund size, $Q_t^{\text{max}}$, may be affected by the lower bound because maximum amount the manager wants to invest could be lower than the minimum required fund size. In these states, the manager may prefer to increase fund size up to the required minimum to ensure that the fund is established. This has a negative impact on investors’ expected return because a larger fund size translates into lower expected excess returns. However, the minimum fund size also has a positive effect on expected returns. This time, in the first stage, the fund manager has to consider the fact that if he fails to raise sufficient capital his payoff will be simply zero. This makes the manager more conservative in choosing the optimal fee.

Note that having a minimum size by itself is not sufficient to generate positive expected returns for investors when managers do not have an incentive to manipulate the beliefs of investors. If there is no information asymmetry regarding a manager’s effort, the manager sets a lower fee to improve the probability that the fund is established, but he still captures all the surplus by increasing the fund size after the realization of uncertainty $\varepsilon_t^P$.

5.4 Performance Fees

In practice, in addition to a 1.5-3 percent management (i.e., “fixed”) fee, most funds charge a 20-25 percent carry interest (i.e., “variable” fee) (Gompers and Lerner, 1999; Phalippou and Gottschalg, 2008; Litvak, 2009). Up to now we have considered a simplified setting where the manager charges only a fixed fee (this is analogous to the framework presented in Berk and Green, 2004). Here, we show that our results carry through when the manager charges a variable fee.

We define the variable fee so that the managers’ period $t$ payoff also includes a component $v_t(P_tW_t - f_t)$ if the realization of the fund’s excess returns is positive, and zero otherwise. In other words, the variable fee is a percent of the fund’s realized return, and has an implicit option-like characteristic that is similar to how simple carry interest is applied in practice. The rest of the model is as before.
We can write the manager’s payoff as

$$\max_{Q_t,e_t} Q_t f_t - C_t + V_{t+1}(E_{t+1}[X|h_t]) + v_t Q_t \int_{\varepsilon(0)} (P_tE_t[W_t] - f_t) dN(\varepsilon^X_t),$$

where $\varepsilon(0)$ is the realization of uncertainty about the manager’s ability which makes the fund’s return equal to zero. Given that fund investors are willing to invest up to the amount that their returns are zero, this optimization problem is subject to:

$$\int_{\varepsilon(0)}^\varepsilon(0) (P_t W_t - f_t) d(\varepsilon^X) + (1 - v_t) \int_{\varepsilon(0)}^\varepsilon(0) (P_t W_t - f_t) dN(\varepsilon^X) \geq 0$$

An immediate observation is that the variable fee cannot capture all the surplus the manager generates unless it is equal to 1, i.e., 100% of returns. However, this is not feasible because when the realized fund returns are negative, investors would lose money, while when fund returns are positive, their return would be zero, ensuring that their participation constraint could not be satisfied.

The variable fee encourages the manager to exert higher effort in equilibrium. Regardless, our results go through. The manager has an incentive to manipulate the beliefs of investors and limits the fund’s size and provides a positive expected excess return to investors, which generates performance persistence over time. We summarize this in the following result.

**Proposition 10** It is not feasible for the manager to capture all the surplus by setting $v^*=1$. For any given $f_t$ and $v_t$, there is a realization of uncertainty $\varepsilon_t^{P,low}$ such that fund investors’ expected excess return $E_t[\alpha_t]\varepsilon_t^P$ is strictly positive for $\varepsilon_t^P > \varepsilon_t^{P,low}$ and zero otherwise. Therefore, at the optimal fee structure $(f_t^*,v_t^*)$ fund investors’ expected return, $E_t[\alpha_t]$, is strictly positive.
5.5 A Simple Model of Matching

Our primary assumption is that entrepreneurs benefit from the value added of the manager and in equilibrium there is positive assortative matching, i.e., managers who have higher perceived ability to add value are more likely to match with better quality firms, where the most value is likely to be created. These are common predictions of models with positive assortative matching (Chemmanur and Fulghieri, 1994; Fernando, Gatchev and Spindt, 2005). Here we provide an informal stylized model of matching to illustrate how these predictions could arise in our context. Assume that at a given period in time entrepreneurs observe managers’ past performance and update their beliefs about managers’ value added. There are two managers, and as a result of observing past performance one of them is perceived as having a higher ability to add value $E[X] = H$ and the other as having a lower ability, $E[X] = L$. Each fund manager has capital $A$ to invest. Managers invest $1$ in each firm to obtains a fixed fraction of each firm (we discuss bargaining below) and there are $N$ firms looking for raising $1$ such that $2A < N$. Each entrepreneur $i$ has a quality $P_i$.

The total value of a firm after a fund manager invests is equal to either $V(P_i, H)$ or $V(P_i, L)$, depending on the quality of the firm and the ability of the manager, with $V(P_i, H) > V(P_i, L)$. The quality of firms is not observed by the fund manager unless the entrepreneur applies for financing and the fund manager screens the firm. At this point an informative signal about the quality of the firm is observed by the fund manager who screened the firm. Fund managers do not share their information with other managers. If a firm does not get financing its payoff is zero. The cost of applying for financing is zero (or low). Therefore, to further simplify the model, assume that all firms simultaneously apply to both fund managers. Managers screen every firm, rank firms based on the observed signals and make offers to the highest quality $A$ firms. Given that matching with a higher ability manager is more beneficial for an entrepreneur, if an entrepreneur gets offers from both managers he accepts the offer from the manager with the higher ability. These are more likely to be better firms given that they are ranked high by both managers. Then managers make offers to the next group of firms and so on. It is obvious that the manager with the lowest ability is
expected to end up with a portfolio of lower quality firms on average because whenever they have overlapping offers his offer is rejected and the pool of potential investments progressively gets worse given that the signals about the quality of the firms is informative. In fact, if the screening is perfect the $A$ highest quality firms will match with the manager who has higher ability and the next $A$ highest quality firms will match with manager who has lower ability to add value.

Our results above rely on the assumption that matching with higher value added manager is more beneficial for entrepreneurs. This will be true unless the fund manager with higher ability to add value captures disproportionately more of the value he generates compared to a manager who has lower ability to add value. In this case the entrepreneurs may prefer to match with the manager who has lower ability to add value. However, this is not likely to be part of a renegotiation proof equilibrium under reasonable bargaining assumptions given that positive assortative matching maximizes social surplus. Matching the entrepreneur with the higher ability manager can always provide higher payoffs to both agents compared to matching the entrepreneur with the lower ability manager. Therefore, one would expect positive assortative matching to be the equilibrium outcome.

6 Conclusion

Anecdotal evidence suggests that many successful private equity funds are oversubscribed. On the other hand, private equity funds appear to generate persistent abnormal returns for their investors, in contrast to mutual funds, which exhibit little or no performance persistence. We argue that private equity funds are fundamentally different from mutual funds because of two reasons: First, private equity funds need to match with good firms, which want to match with managers who have higher ability to add value. Second, there is greater asymmetry of information regarding private equity fund managers’ ability to add value because observed returns could be a function of selection in addition to value adding ability. Therefore, there is a high incentive for private equity fund managers to attempt to manipu-
late the beliefs of firms about their ability. In particular, by exerting effort to select better firms a manager tries to improve firms’ beliefs about his ability to add value. Managers also keep fund size small because it is less costly to improve the quality of firms in a smaller portfolio.

In the signal jamming equilibrium we develop, firms are not fooled and they correctly form an unbiased expectation. Nevertheless, managers try to manipulate beliefs of entrepreneurs limit the sizes of their fund because otherwise their probability of matching with good firms would decrease. Our model not only explains differences in performance persistence between mutual and private equity funds but also provides new predictions about how our results would vary with the cross sectional differences among funds and over time.
Appendix

Proof of Proposition 1: The Lagrangian for the fund manager’s optimization problem at time $t$, for a given $f_t > 0$, is

$$\max_{Q_t, e_t} \mathcal{L}^2 = Q_t f_t - C_t + \lambda_t (P_t E_t[W_t] - f_t).$$

The Kuhn-Tucker conditions are:

$$f_t - \frac{\partial C}{\partial Q_t} - \lambda_t P_t \frac{\partial S}{\partial Q_t} = 0, \quad (9)$$

$$-\frac{\partial C}{\partial e_t} + \lambda_t E[W_t] \frac{\partial P}{\partial e_t} = 0, \quad (10)$$

$$\lambda_t (P_t E[W_t] - f_t) = 0. \quad (11)$$

Consider the case that $\lambda_t = 0$. In this case the FOC with respect to effort, (10), can only be satisfied if effort is equal to zero. However, the FOC with respect to quantity, (9), cannot be satisfied for $f_t > 0$ given that $\frac{\partial C}{\partial Q_t} = 0$ when effort is equal to zero. Therefore, the solution must have $\lambda_t > 0$, so that $P_t E[W_t] - f_t = 0$.

There may also be a region where the participation constraint of investors cannot be satisfied for very low realizations of $\varepsilon^T_t$. In that case the fund cannot be established in period $t$ and investor returns are again equal to zero. □

Proof of Proposition 2: The Lagrangian for the fund manager’s optimization problem with respect to the fee is

$$\max_{f_t} \mathcal{L}^1 = E[Q_t f_t - C_t],$$

subject to $Q_t$ and $e_t$ being defined from (9) and (10), respectively, as well as investors’ participation constraint. Note that, from Proposition 1, it may not always be possible to satisfy investors’ participation constraint, so that the region where a fund cannot be
established is a function of $f_t$. Using Leibniz’s rule, the FOC can be written as

$$
\frac{d\mathcal{L}}{df_t} = -\frac{\partial \varepsilon^P_t}{\partial f} (Q_t f_t - C_t)|\varepsilon^P_t + \int_{\varepsilon^P_t} \left( \frac{\partial [Q_t f_t - C_t]}{\partial f_t} + \frac{\partial [Q_t f_t - C_t]}{\partial Q_t} dQ_t + \frac{\partial [Q_t f_t - C_t]}{\partial e_t} de_t \right) dN(\varepsilon^P) = 0.
$$

When the fund is not established, i.e., $\varepsilon^P_t < \varepsilon^P_t$, $Q_t f_t - C_t$ has to equal zero, as otherwise the manager could create slack in the participation constraint of investors either by decreasing fund size or increasing effort. In the region where the fund is established we know that

$$
\frac{\partial [Q_t f_t - C_t]}{\partial Q_t} = \frac{\partial \mathcal{L}_2}{\partial Q_t} = \lambda_t P_t \frac{\partial S}{\partial Q_t} + \lambda_t E[W_{t+1} \mid h_t] \frac{\partial E[P \mid \varepsilon^P_t]}{\partial e_t} = 0
$$

from the first order conditions.

Proof of Propositions 3 and 4: From (5) and (6), we see that for a given $f_t$, the Lagrangian for the fund manager’s problem is

$$
\mathcal{L}^2 = Q_t f_t - C_t + V_{t+1}(E_{t+1}[X \mid h_t]) + \lambda_t (P_t E[W_t] - f_t).
$$

The first order conditions are

$$
f_t - \frac{\partial C}{\partial Q_t} - \lambda_t P_t \frac{\partial S}{\partial Q_t} = 0,
$$

□
\[-\frac{\partial C}{\partial e_t} + \frac{\partial V_{t+1}}{\partial E_{t+1}[X|h_t]} \frac{\partial E_{t+1}[X|h_t]}{\partial e_t} = 0, \]  \hspace{1cm} (14)

\[\lambda_t(P_tE[W_t] - f_t) = 0. \hspace{1cm} (15)\]

Note that since effort is not observed, the derivative of the participation constraint with respect to effort is zero. Effort is determined by \(\frac{\partial C}{\partial e_t} = \frac{\partial V_{t+1}}{\partial E_{t+1}[X|h_t]} \frac{\partial E_{t+1}[X|h_t]}{\partial e_t}\), with the manager balancing the cost of effort against its benefit in improving his payoff in the future. Therefore, the manager always exerts some effort in order to manipulate investors’ beliefs. A quick check shows that \(\lambda_t\) can be equal to zero for certain parameter values. When \(\lambda_t = 0\), fund size \(Q_t\) is determined by \(f_t = \frac{\partial C}{\partial Q_t}\), so that the manager equalizes the marginal cost of size (at the conjectured effort level) to the fixed fee. This is the largest fund size that the manager is willing to operate because when \(\lambda_t > 0\) the manager selects a smaller \(Q_t\) compared to the case when \(\lambda_t = 0\) given that \(P_t\frac{\partial S}{\partial Q_t} > 0\). Define \(Q_{t}^{\text{max}}\) as this maximal fund size, defined by \(f_t = \frac{\partial C}{\partial Q_t}\). Define as well \(e_t^{\text{min}}\) as the equilibrium level of effort at the maximal fund size \(Q_{t}^{\text{max}}\). Note that \(e_t^{\text{min}}\) also defines a lower bound on the level of effort the fund manager will find it optimal to exert since for values of \(Q_t < Q_{t}^{\text{max}}\), the optimal level of effort will be no less than \(e_t^{\text{min}}\) from (14). This establishes Proposition 3.

Note now that investors’ expected return, \(E_t[\alpha_t|x_t^P]\), will be positive when the realization of \(x_t^P\) is greater than \(x_t^{P,low} \equiv \frac{f_t}{E[W_t(Q_{t}^{\text{max}})]} - P_t(e_t^{\text{min}})\), and zero otherwise. For every \(f_t\) there is a realization of \(x_t^P\) such that investors’ expected return is greater than zero.

Note finally that if \(x_t^P < x_t^{P,low}\), there are two possibilities. First, if investors’ participation constraint binds, i.e., if \(P_tE[W_t] - f_t = 0\), for \(Q_t < Q_{t}^{\text{max}}\), this implies that \(f_t > \frac{\partial C}{\partial Q_t}\) and therefore that \(\lambda_t > 0\). The other possibility is that for very low realizations of \(x_t^P\), it will not be possible to satisfy the participation constraint of investors even as fund size approaches zero. In this case the fund will not be established in that period. We define \(\xi_t^P\) as the minimum realization of \(x_t^P\) such that a fund is established in period \(t\). This completes the proof for Proposition 4. \(\Box\)
Proof of Proposition 5: The Lagrangian for the fund manager’s optimization problem with respect to the fee \( f_t \) is

\[
\max_{f_t} L^1 = E[Q_t f_t - C_t + V_{t+1}(E_{t+1}[X|h_t])],
\]

subject to \( Q_t \) and \( e_t \) being defined from (13) and (14), respectively, as well as investors’ participation constraint. Similar to the proof of Proposition 2, the FOC can be written as:

\[
\frac{dL^1}{df_t} = -\frac{\partial L^1}{\partial P_t} (Q_t f_t - C_t + V_{t+1}(E_{t+1}[X|h_t]))|\varepsilon_t^P
\]

\[
+ \int_{\varepsilon_t^P} \left( \frac{\partial [Q_t f_t - C_t + V_{t+1}(E_{t+1}[X|h_t])]}{\partial f_t} dQ_t + \frac{\partial [Q_t f_t - C_t + V_{t+1}(E_{t+1}[X|h_t])]}{\partial e_t} dS_t \right) dN(\varepsilon_t^P) = 0.
\]

However, as in the proof of Proposition 2, we know that \( (Q_t f_t - C_t + V_{t+1}(E_{t+1}[X|h_t]))|\varepsilon_t^P \) must be equal to zero, and

\[
\frac{\partial [Q_t f_t - C_t + V_{t+1}(E_{t+1}[X|h_t])]}{\partial Q_t} = \lambda_t P_t C_t + \lambda_t P_t \frac{\partial S}{\partial Q_t},
\]

since \( \frac{\partial C^2}{\partial Q_t} = 0 \) from the envelope theorem. Moreover,

\[
\frac{\partial [Q_t f_t - C_t + V_{t+1}(E_{t+1}[X|h_t])]}{\partial e_t} = \frac{\partial C^2}{\partial e_t} = 0.
\]

We can therefore rewrite the FOC as

\[
\int_{\varepsilon_t^P} \left( Q_t + \lambda_t P_t \frac{\partial S}{\partial Q_t} \right) dN(\varepsilon_t^P) = 0.
\]

(16)

We can see that \( E[\lambda_t \frac{dQ_t}{df_t}] < 0 \) because from the second stage problem we know that when investors’ participation constraint binds \( \lambda_t \frac{dQ_t}{df_t} < 0 \), and otherwise \( \lambda_t \frac{dQ_t}{df_t} = 0 \). Therefore (16) defines a solution for the optimal fee, \( f_t^A \), such that the manager trades off higher fees against a lower expected fund size and a lower probability of establishing the fund. \( \Box \)

Proofs of Corollary 6: Here we provide the proof only with respect to change in \( w \). The proofs of all other remaining results in Corollary 6 are similar as they rely on the same intuition and are therefore omitted. It is straightforward to show that optimal effort \( e_t \) decreases as \( w_t \) increases. To see this, note that \( \frac{\partial E_{t+1}[X|h_t]}{\partial e_t} \) is decreasing in \( w_t \) from the definition of \( E_{t+1}[X|h_t] \) in (4). Therefore, for every \( Q_t \) and \( f_t \), the FOC for effort, (14), will yield a lower value for \( e_t \). For fund size \( Q_t \), note that for \( \lambda_t = 0 \), the equilibrium value of
obtained from (13) is decreasing in \( e \) for every \( f_t \). Put together, this implies that \( Q_t^{\text{max}} \) increases and \( e_t^{\text{min}} \) decreases as \( w_t \) increases. This implies that, for every \( f_t \), the cutoff value \( \varepsilon_t^{P \text{low}} = \frac{\int E[W_t(Q_t^{\text{max}})]}{E[W_t(Q_t^{\text{max}})]} - P_t(e_t^{\text{min}}) \) must increase as \( w_t \) increases.

We now characterize what happens to the fee \( f_t \). To find the sign of \( \frac{d f_t}{d w_t} \), we take the total derivative of \( \partial L_1 \partial f_t \) with respect to \( w_t \):

\[
\frac{\partial^2 L_1}{\partial f_t^2} \frac{df_t}{dw_t} + \frac{\partial^2 L_1}{\partial f_t \partial w_t} = 0.
\]

Since \( \frac{\partial^2 L_1}{\partial f_t^2} < 0 \) from the SOC, we have

\[
\text{sign} \left( \frac{df_t}{dw_t} \right) = \text{sign} \left( \frac{\partial^2 L_1}{\partial f_t \partial w_t} \right).
\]

This latter expression can be written as

\[
\frac{\partial^2 L_1}{\partial f_t \partial w_t} = \frac{\partial \varepsilon_t^P}{\partial w_t} \left( Q_t + \lambda_t P_t \frac{\partial S}{\partial Q_t} dQ_t \right) \varepsilon_t^P + \int \frac{\partial}{\varepsilon_t^P} dw_t \left( Q_t + \lambda_t P_t \frac{\partial S}{\partial Q_t} dQ_t \right) dN(\varepsilon_t^P).
\]

Note that \( \frac{\partial \varepsilon_t^P}{\partial w_t} > 0 \), which can be seen by solving for \( \varepsilon_t^P \) from investors’ participation constraint: higher \( w_t \) implies lower effort, which results in higher \( \varepsilon_t^P \). On the other hand, the term \( (Q_t + \lambda_t P_t \frac{\partial S}{\partial Q_t} dQ_t)\varepsilon_t^P \) is negative since for large realizations of \( \varepsilon_t^P \), investors’ participation constraint does not bind, so that \( \lambda_t = 0 \) and \( Q_t \) is positive. Therefore, for (16) to be satisfied for lower realizations of \( \varepsilon_t^P \) it must be that \( (Q_t + \lambda_t P_t \frac{\partial S}{\partial Q_t} dQ_t) dN(\varepsilon_t^P) < 0 \). As a result, \( -\frac{\partial \varepsilon_t^P}{\partial w_t} (Q_t + \lambda_t P_t \frac{\partial S}{\partial Q_t} dQ_t) \varepsilon_t^P > 0 \).

For the second term, note that \( \frac{\partial Q_t}{\partial w_t} > 0 \) for every \( \varepsilon_t^P \), so that \( \int \varepsilon_t^P \frac{\partial Q_t}{\partial w_t} dN(\varepsilon_t^P) > 0 \). Consider now \( \frac{\partial}{\partial w_t} \left( \lambda_t P_t \frac{\partial S}{\partial Q_t} dQ_t \right) = \lambda_t P_t \frac{\partial S}{\partial Q_t} \frac{\partial}{\partial w_t} \left( \frac{dQ_t}{df_t} \right) \) since neither \( \lambda_t \), \( P_t \), nor \( \frac{\partial S}{\partial Q_t} \) are direct functions of \( w_t \). Note as well that in the region where \( \lambda_t > 0 \), \( \frac{dQ_t}{df_t} \) is obtained from investors’ participation constraint, and is can be obtained from totally differentiating the participation constraint:

\[
P_t \frac{\partial E[W_t]}{\partial Q_t} dQ_t - df_t = 0 \Rightarrow \frac{dQ_t}{df_t} = 1 / P_t \frac{\partial E[W_t]}{\partial Q_t} < 0.
\]
since $\frac{\partial E[W_t]}{\partial Q_t} < 0$. Note, however, that neither $P_t$ nor $\frac{\partial E[W_t]}{\partial Q_t}$ are direct functions of $w_t$, so that
\[ \frac{\partial}{\partial w_t} \left( \lambda_t P_t \frac{\partial S}{\partial Q_t} \frac{dQ_t}{df_t} \right) = 0. \]
We can therefore conclude that
\[ \frac{\partial^2 L^1}{\partial f_t \partial w_t} = -\frac{\partial \xi_t^P}{\partial w_t} \left( Q_t + \lambda_t P_t \frac{\partial S}{\partial Q_t} \frac{dQ_t}{df_t} \right) |\xi_t^P| + \int_{\xi_t^P} \frac{\partial Q_t}{\partial w_t} dN(\xi_t^P) > 0. \]
Hence, $\frac{df}{dw_t} > 0$ and a fund manager will choose higher fees when $w_t$ is larger. Therefore, $\xi_t^{P,low}$ must be increasing in $w_t$ and the probability of fund investors having positive expected return decreases. □

**Proof of Proposition 8:** The objective function is
\[ V_t(E_t[X|h_{t-1}]) = \max_{Q_t, o_t} [Q_tf_t - K(Q_t, o_t) + V_{t+1}(E_{t+1}[X|h_t])]. \]
The Kuhn Tucker conditions are:
\[ f_t - \frac{\partial K}{\partial Q_t} - \lambda(P_t) \frac{\partial S}{\partial Q_t} = 0, \]
\[ -\frac{\partial K}{\partial o_t} + \frac{\partial V_{t+1}}{\partial E[X]} \frac{\partial E[X]}{\partial o_t} = 0, \]
\[ \lambda(P_tE[W_t] - f_t) = 0. \]
Note that these conditions are almost identical to the ones in Proposition 3. A similar proof therefore applies. □

**Proof of Proposition 9:** The manager determines $Q_t$ and $e_t$ considering the minimum size limit:
\[ \max_{Q_t, e_t} L^2 = Q_tf_t - C_t + \lambda_t(P_tE_t[W_t] - f_t) + \beta_t(Q_t - Q_{min}), \]
where $\beta_t$ is the Lagrange multiplier on the constraint that $Q_t \geq Q_{\min}$. The FOCs are

\[
f_t - \frac{\partial C}{\partial Q_t} - \lambda_t f_t \frac{\partial S}{\partial Q_t} + \beta_t = 0, \tag{17}
\]

\[
- \frac{\partial C}{\partial e_t} + \frac{\partial V_{t+1}}{\partial E_{t+1}} \frac{\partial E_{t+1} [X|h_t]}{\partial e_t} = 0, \tag{18}
\]

\[
\lambda_t (P_t E_t [W] - f_t) = 0, \tag{19}
\]

\[
\beta_t (Q_t - Q_{\min}) = 0. \tag{20}
\]

When there is no exogenously given minimum fund size $Q_{\min}$ and $\lambda_t = 0$, fund size is determined by $f_t = \frac{\partial C}{\partial Q_t}$. However, if this yields a fund size lower than $Q_{\min}$, i.e., $\beta_t > 0$, then $f_t + \beta_t = \frac{\partial C}{\partial Q_t}$ determines fund size. In other words, the manager increases the size of the fund to ensure the participation of investors when the optimal quantity is below $Q_{\min}$.

We can define the new maximum quantity with the minimum fund size limit as $Q_{\max}^{\min} = \max\{Q_{\max}^t, Q_{\min}\}$. Investors’ expected return $E_t[\alpha_t]$ is positive when the realization of $\varepsilon_t^P > \frac{b}{E[W(Q_{\min})]} - P(\varepsilon_{\min}) = \varepsilon_t^{P,low}$, and zero otherwise. For every $f_t$ there is a realization of $\varepsilon_t^P$ such that investors’ expected return is larger than zero. Note that, when $Q_{\max}^{\min} = Q_{\min}$, this decreases the expected return for investors since it requires the fund manager to raise more funds that are optimal. The Lagrangian for the fund manager’s optimization problem with respect to the fee is:

\[
\max_{f_t} L^1 = E \left[ Q_t f_t - C_t + V_{t+1} \left( E_{t+1} [X|h_t] \right) \right].
\]

Similar to proof of Proposition 2, the FOC can be written as:

\[
\frac{dL^1}{df_t} = - \frac{\partial \varepsilon_t^P}{\partial f_t} (Q_t f_t - C_t) |\varepsilon_t^P| + \int_{\varepsilon_t^P} \left( \frac{\partial [Q_t f_t - C_t]}{\partial f_t} + \frac{\partial [Q_t f_t - C_t]}{\partial Q_t} \frac{dQ_t}{df_t} \right) dN(\varepsilon_t^P) = 0.
\]

However, as in the the proof of Proposition 2, we know that $(Q_t f_t - C_t) |\varepsilon_t^P = 0$ and $\frac{\partial [Q_t f_t - C_t]}{\partial Q_t} = \ldots$
\[ \frac{\partial^2 C}{\partial Q_t^2} + \lambda_t P_t \frac{\partial S}{\partial Q_t} - \beta_t = \lambda_t P_t \frac{\partial S}{\partial Q_t} - \beta_t \] since \( \frac{\partial S}{\partial Q_t} = 0 \) from the envelope theorem. Similarly \( \frac{\partial \left[ Q_t f_t - C_t + V_{t+1}(E_{t+1}[X|h_t]) \right]}{\partial e_t} = \frac{\partial^2 C}{\partial e_t^2} = 0. \) We can therefore rewrite the FOC as

\[ \int_{\varepsilon_t^P} \left( Q_t + \lambda_t P_t \frac{\partial S}{\partial Q_t} \frac{dQ_t}{df_t} - \beta_t \right) dN(\varepsilon_t^P) = 0. \] (21)

The optimal fee, \( f_t^0 \), trades off larger fees (first term) with a lower fund size (second term) and the shadow cost of the fund not being established. □

**Proof of Proposition 10:** Define \( \varepsilon_t^{X,0} \) as the realization of \( \varepsilon_t^X \) that equalizes realized return of fund investors to zero. Analysis of the participation constraint reveals that \( v_t \) must be strictly less than 1, as otherwise investors’ participation constraint could not be satisfied.

The Kuhn Tucker conditions are:

\[

f_t - \frac{\partial C_t}{\partial Q_t} + v_t \int_{\varepsilon_t^{X,0}} (P_t E_t[W_t] - f_t) dN(\varepsilon_t^X) + Q_t v_t \frac{\partial}{\partial Q_t} \left( \int_{\varepsilon_t^{X,0}} (P_t E_t[W_t] - f_t) dN(\varepsilon_t^X) \right) \\
+ \lambda_t \frac{\partial}{\partial Q_t} \left( \int_{\varepsilon_t^{X,0}} (P_t W_t - f_t) d(\varepsilon_t^X) + (1 - v_t) \int_{\varepsilon_t^{X,0}} (P_t W_t - f_t) dN(\varepsilon_t^X) \right) = 0,
\]

\[

- \frac{\partial C_t}{\partial e_t} + \frac{\partial V_t}{\partial X_t} \frac{\partial X_t}{\partial e_t} + v_t Q_t \int_{\varepsilon_t^{X,0}} \frac{\partial P_t}{\partial e_t} W_t dN(\varepsilon_t^X) + \lambda_t \frac{\partial}{\partial e_t} \left( \int_{\varepsilon_t^{X,0}} (P_t W_t - f_t) dN(\varepsilon_t^X) \\
+ (1 - v_t) \int_{\varepsilon_t^{X,0}} (P_t W_t - f_t) dN(\varepsilon_t^X) \right) = 0,
\]

\[

\lambda_t \left( \int_{\varepsilon_t^{X,0}} (P_t W_t - f_t) dN(\varepsilon_t^X) + (1 - v_t) \int_{\varepsilon_t^{X,0}} (P_t W_t - f_t) dN(\varepsilon_t^X) \right) = 0.
\]

Note that \( \lambda_t = 0 \) can be a potential solution. When \( \lambda_t = 0 \), the maximum quantity \( Q_t^{\text{max}} \) is determined by \( f_t - \frac{\partial C_t}{\partial Q_t} + v_t \int_{\varepsilon(0)} (P_t E_t[W_t] - f_t) d(\varepsilon_t^X) + Q_t v_t \frac{\partial}{\partial Q_t} \left( \int_{\varepsilon(0)} (P_t E_t[W_t] - f_t) d(\varepsilon_t^X) \right) \) while the \( e_{\text{min}} \) is determined by \( -\frac{\partial C_t}{\partial e_t} + \frac{\partial V_t}{\partial X_t} \frac{\partial X_t}{\partial e_t} + v_t Q_t \int_{\varepsilon(0)} \frac{\partial P_t}{\partial e_t} W_t d(\varepsilon_t^X) \). As before by placing \( Q^{\text{max}} \) and \( e_{\text{min}} \) in the investors’ participation constraint one can derive the minimum \( \varepsilon_t^{P,\text{low}} \) for any \( f_t \) and \( v_t \) such that investors’ expected excess return is positive. If this is true for any \( f_t \) and \( v_t \), it will be true for the optimal fees \( f_t^* \) and \( v_t^* \) such that \( f_t^* < \infty \) and \( v_t^* < 1 \). We
have argued above that \( v^*_i < 1 \) and \( f^*_i \) is bounded from the first stage optimization problem because, as before, larger \( f_i \) imply smaller fund size and lower probability of establishing the fund. \( \square \)
References


Phalippou, Ludovic, 2009, “Venture Capital Funds: Performance Persistence and Flow-
Performance Relation,” forthcoming *Journal of Banking and Finance*.


