# Illiquid Asset Investing

This version: 01-13-2013

## Abstract

After taking into account biases induced by infrequent trading and selection, it is unlikely that illiquid asset classes have higher risk-adjusted returns than traditional liquid stock and bond markets. On the other hand, there are significant illiquidity premiums within asset classes. Portfolio choice models incorporating illiquidity risk recommend only modest holdings of illiquid assets. Investors should demand high risk premiums for investing in illiquid assets.

## 1. Liquidating Harvard

No-one thought it could happen to Harvard.[1]

In 2008, Harvard University's endowment – the world's largest – fell victim to the world-wide plunge in asset prices triggered by the financial crisis. In contrast to its 15% average annual returns since 1980, Harvard's endowment suffered its worst decline in history, falling 22% in value between July 1 and October 31, 2008. More than $8 billion in value had been wiped from Harvard's endowment in three months.

Concerned with the impending budget shortfall due to the collapse in the endowment, University President Drew Faust and Executive Vice President Edward Forst sounded the alarm by sending a memo to the Council of Deans on December 2, 2008. They asked each school to cut expenses and compensation, and to scale back ambitions in the face of reduced revenue. As bad as the reported losses were, they cautioned that the true losses were even worse: "Yet even the sobering

[1] This is based on "Liquidating Harvard," Columbia CaseWorks ID #100312.

figure is unlikely to capture the full extent of actual losses for this period, because it does not reflect fully updated valuations in certain managed asset classes, most notably private equity and real estate."[2]

Harvard relied on endowment earnings to meet a large share of university expenses. In its fiscal year ending June 30, 2008, more than one-third of operating revenue came from endowment income. For some of the university's individual departments, the proportion was even higher: the Radcliffe Institute for Advanced Study derived 83% of its revenue from the endowment, the Divinity School 71%, and the Faculty of Arts and Sciences 52%.

Harvard Management Company (HMC), the funds manager of Harvard's endowment, was one of the early adopters of the *endowment model*, which recommends that long-term investors should hold lots of illiquid, alternative assets, especially private equity and hedge funds. Advocated by David Swensen in his influential book, *Pioneering Portfolio Management*, the endowment model was based on the economic concept of diversification originally attributable to Harry Markowitz (1952, see Chapter XX). Through diversification, a portfolio of many low-correlated assets has a risk-return trade-off superior to that of conventional portfolios consisting of only stocks and bonds. Swensen went further and advocated holding large proportions of illiquid private equity and hedge funds. Not only were these assets supposed to have low correlations to stocks and bonds, but they potentially carried an illiquidity risk premium.

Swensen argued that in liquid markets, the potential for making excess returns (or "alpha," see Chapter XX) was limited. In these markets, crowded with thousands of active managers vying for an edge, information was freely available and almost everyone has access to it. Illiquid asset

[2] Financial Update to the Council of Deans, December 2, 2008 from Faust and Forst.

markets, like venture capital and private equity, had large potential payoffs for investors who had superior research and management skills. Swensen argued that alpha was not competed away in illiquid assets because most managers have short horizons. University endowments, with their longer horizons, would seem to have an advantage in illiquid assets. Swensen recommended that long-term institutions with sufficient resources who can carefully select expert managers in alternative, illiquid assets could achieve superior risk-adjusted returns.

Dutifully following Swensen's advice, many endowments, including Harvard, loaded up with illiquid assets during the 1990s. In 2008, HMC held 55% of its portfolio in hedge funds, private equity, and real assets. Only 30% was in developed-world equities and fixed income, with the remainder of its portfolio in emerging market equities and high-yield bonds.

In its desperate need for cash, HMC tried to sell some of its $1.5 billion private equity portfolio, which included marquee names such as Apollo Investment and Bain Capital. But buyers in secondary markets demanded huge discounts. Nina Munk, a journalist writing in *Vanity Fair*, recounts a surreal conversation between the CIO of HMC, Jane Mendillo, and a money manager specializing in alternative investments:[3]

> Funds Manager: Hey look, I'll buy it back from you. I'll buy my interest back.
>
> Mendillo: Great
>
> Funds Manager: Here, I think its worth you know, today the [book] value is a dollar, so I'll pay you 50 cents.
>
> Mendillo: Then why would I sell it?

[3] Nina Munk, "Richard Harvard, Poor Harvard," Vanity Fair, August 2009.

> Funds Manager: Well, why are you? I don't know. You're the one who wants to sell, not me. If you guys want to sell, I'm happy to rip your lungs out. If you are desperate, I'm a buyer.
>
> Mendillo: Well, we're not desperate.

But in truth Harvard was desperate.

The reaction to Faust and Forst's cost-cutting memo was swift and sharp. Faculty, students, and alumni were incredulous. Alan Dershowitz, a famous professor at Harvard Law School, said:[4] "Apparently nobody in our financial office has read the story in Genesis about Joseph interpreting Pharaoh's dream… You know, during the seven good years you save for the seven lean years."

All the short-term decisions for Harvard leaders and Mendillo at HMC were painful: slashing budgets, hiring freezes, and the postponement of the university's planned Allston science complex. Asset-liability management for Harvard University had failed. In the longer term, was the endowment model with illiquid, alternative assets still appropriate?

# 2 Illiquid Asset Markets

## 2.1 Sources of Illiquidity

Vayanos and Wang (2012) provide a taxonomy of how illiquidity arises due to market imperfections:

1. Clientele effects and participation costs

[4] Quoted by Munk, N., Rich Harvard, Poor Harvard, Vanity Fair, August 2009.

Entering markets can be costly; investors often must spend money, time or energy to learn their way around and gain the necessary skills. In many large, illiquid asset markets, only certain types of investors with sufficient capital, expertise, and experience can transact.

2. Transactions costs

   These include commissions, taxes, and for certain illiquid assets, the costs of due diligence, title transfers and the like, as well as the bread and butter costs incurred for trading. It also includes fees paid to lawyers, accountants, and investment bankers. Academics sometimes assume that investors can trade whenever they want as long as they pay (sometimes a substantial) cost, but this is not always true because of…

3. Search frictions

   For many assets, you need to search to find an appropriate buyer or seller. Only certain investors have the skills to value a complicated structured credit product. There are few large investors with sufficient capital to invest in skyscrapers in major metropolitan areas. You might have to wait a long time to transact.

4. Asymmetric information

   Markets can be illiquid because one investor has superior knowledge compared with other investors. Fearing they'll be fleeced, investors become reluctant to trade. When asymmetric information is extreme – all the products are lemons, and no-one wants to buy a lemon – markets break down. Many liquidity freezes are caused by these situations. The presence of asymmetric information also causes investors to look for non-predatory counterparties, so information is a form of search friction.

5. Price impact

Large trades will move markets.

6. Funding constraints

   Many of the investment vehicles used to invest in illiquid assets are highly leveraged. Even investing in a house requires substantial leverage for most consumers. If access to credit is impaired, investors cannot transact in illiquid asset markets.

Illiquid asset markets are characterized by many, and sometimes all, of the market imperfections on this list. I refer to these effects as "illiquidity." On the basis of this reasoning, all assets are at least somewhat illiquid – even the large-cap equities that trade several times every second – but of course some assets are much more illiquid than others. Illiquidity manifests as infrequent trading, small amounts being traded, and low turnover. Some of the intervals between trades in illiquid markets can extend to decades. Table 1, adapted from Ang, Papanikolaou and Westerfield (2012), lists average intervals between trading and turnover for several asset classes.[5] First, note that…

[Table 1 here]

## 2.2 Most Asset Classes are Illiquid

Except for "plain vanilla" public equities and fixed income, most asset markets are characterized by long periods, sometimes spanning decades, between trades, and they have very low turnover. Even within the equities and fixed income asset classes, some sub-asset classes are highly illiquid. Equities trading in pink-sheet over-the-counter (OTC) markets may go for a week without trading. The average municipal bond trades only twice per year, and the entire muni bond market has an annual turnover of less than 10% (see also Chapter XX). In real estate

[5] See Ang, Papanikolaou and Westerfield (2012) for additional references behind the numbers in Table 1 and other references in this section.

markets, the typical holding period is 4 to 5 years for single-family homes and 8 to 11 years for institutional properties. Holding periods for institutional infrastructure can be 50 years or longer, and works of art sell every 40-70 years, on average. Thus, most asset markets are illiquid in the sense that they trade infrequently and turnover is low.

### 2.3 Illiquid Asset Markets are Large

The illiquid asset classes are large and rival the size of the public equity market. In 2012, the market capitalization of the NYSE and NASDAQ was approximately $17 trillion. The estimated size of the U.S. residential real estate market is $16 trillion, and the direct institutional real estate market is $9 trillion. In fact, the traditional public, liquid markets of stocks and bonds are smaller than the total wealth held in illiquid assets.

### 2.4 Investors Hold Large Amounts of Illiquid Assets

Illiquid assets dominate most investors' portfolios. For individuals, illiquid assets represent 90% of their total wealth, which is mostly tied up in their house – and this is before counting the largest and least liquid component of individuals' wealth, human capital (see Chapter XX). There are high proportions of illiquid assets in rich investors' portfolios, too. High net worth individuals in the U.S. allocate 10% of their portfolios to "treasure" assets like fine art and jewelry. This rises to 20% for high net worth individuals in other countries.[6]

The share of illiquid assets in institutional portfolios has increased dramatically over the last 20 years. The National Association of College and University Business Officers (NACUBO) reported that, in 2011, the average endowment held a portfolio weight of more than 25% in alternative assets versus roughly 5% in the early 1990s. A similar trend is evident among pension

[6] See Profit or Pleasure? Exploring the Motivations Behind Treasure Trends, Wealth Insights, Barclays Wealth and Investment Management, 2012.

funds. In 1995, they held less than 5% of their portfolios in illiquid alternatives, but today the figure is close to 20%.[7]

## 2.5 Liquidity Dries Up

Many normally liquid asset markets periodically become illiquid. During the 2008-2009 financial crisis, the market for commercial paper (or the "money market") – usually very liquid – experienced "buyers' strikes" by investors unwilling to trade at any price. This was not the first time that the money market froze: trading in commercial paper also ceased when the Penn Central railroad collapsed in 1970. In both cases, the money market needed to be resuscitated by the Federal Reserve, which stepped in to restore liquidity.

During the financial crisis, illiquidity also dried up in the repo market (which allows investors to short bonds), residential and commercial mortgage-backed securities, structured credit, and the auction rate security market (a market for floating rate municipal bonds, see Chapter XX ). The last example was one of the first markets to become illiquid at the onset of the financial crisis in 2008 and at the time of writing in 2013 is still frozen. This market is dead in its present form.

Illiquidity crises occur regularly because liquidity tends to dry up during periods of severe market distress. The Latin American debt crisis in the 1980s, the Asian emerging market crisis in the 1990s, the Russian default crisis in 1998, and of course the financial crisis of 2008-2009 were all characterized by sharply reduced liquidity, and in some cases liquidity completed evaporated in some markets. Major illiquidity crises have occurred at least once every 10 years, most in tandem with large downturns in asset markets.

[7] See Global Pension Asset Study, Towers Watson, 2011.

### 2.6 Summary

Illiquid asset classes taken as a whole are larger than the traditional liquid, public markets of stocks and bonds. Even normally liquid markets periodically become illiquid. Most investors' wealth is tied up in illiquid assets. Thus, asset owners must consider illiquidity risk in the construction of their portfolios. Doing this requires estimating risk-return trade-offs of illiquid assets, but measuring illiquid asset returns is not straightforward.

## 3 Illiquid Asset Reported Returns are Not Returns

As Faust and Forst note in their memo to Harvard's Council of Deans, the true illiquid asset losses were greater than the reported ones, which leads us to an important corollary. *Reported illiquid asset returns are not returns*. With respect to these assets, three key biases cause people to overstate expected returns and understate risk:

1. Survivorship bias
2. Infrequent sampling
3. Selection bias

In illiquid asset markets, investors must be highly skeptical of reported returns.

### 3.1 Survivorship Bias

Survivorship bias results from the tendency of poorly performing funds to stop reporting. Many of these funds ultimately fail – but we only rarely count their failures. This makes true illiquid asset returns worse than the reported data.

Here's an analogy: Suppose we wanted to test the hypothesis that smoking is bad for you. We're only going to run our tests on a sample of smokers that have puffed cigarettes for at least 20 years and are in good health today. Lo and behold, we conclude that this *select* group of smokers has a slightly better mortality rate than the general population. From this analysis, can we conclude that smoking is actually good for you? Of course not! We have taken a *biased sample* that contains smokers blessed with longevity who are, so far, invulnerable to the detrimental effects of smoking. If you were to take up smoking today what are the odds that you would end up in this lucky group 20 years later? Or would you die from emphysema (or heart disease, or lung cancer, etc.) before the experiment could be repeated in 20 years time?

Surviving funds in illiquid asset management are like those lucky, long-lived smokers. We observe the returns of surviving funds precisely because they are still around, and they are generally above average. All of the unlucky illiquid managers disappear and thus stop reporting returns. These non-survivors have below average returns.[8] Industry analysis of buy-out funds, venture capital funds, or [insert your favorite illiquid asset class] tends to encompass only firms that have survived over the period of the analysis. But do we know that the small venture capital firm we're investing in today will be around 10 years later? Existing firms and funds, by dint of being alive today, tend to have better-than-average track records. This produces reported returns of these asset classes that are too good to be true.

The only way to completely remove the effect of survivorship bias is to observe the entire population of funds. *In illiquid asset markets we never observe the full universe.*

We can gauge the impact of survivorship bias with mutual funds, which are required to report their returns to the SEC because they fall under the 1940 Investment Act. This allows us to see the whole mutual fund universe (at least when the funds become registered) and to compute the effect of survivorship bias. (I provide more details in Chapter XX.) Survivorship bias knocks at

[8] Jorion and Goetzmann (1999) argue that survivorship bias partly explains the high equity premium, see Chapter XX: Countries where we have long histories of equity returns are, by definition, those countries where equity investments have prospered.

least 1-2% off the estimates of expected returns of mutual funds, and the survivorship effect can go up to 4% in some cases. Take these as lower bounds for illiquid asset managers. Chapters XX and XX covering hedge funds and private equity, respectively, show that managers of these investment vehicles often further massage (or manipulate) returns because standardized disclosure is not required and the underlying asset values are not readily observable. In Chapter XX, I show the effect of survivorship and reporting biases for hedge funds is even larger than for mutual funds.

There are data biases other than survivorship bias: for funds specializing in very illiquid assets, reporting returns to database vendors is almost always voluntary. This introduces *reporting biases*.[9] Survivorship bias results when your fund is in the database now and you stop reporting returns because you know your returns are going to be low. Reporting bias is that you don't start reporting your returns in the first place because your fund never achieves a sufficiently attractive track record.

### 3.2 Infrequent Trading

With infrequent trading, estimates of risk – volatilities, correlations, and betas – are too low when computed using reported returns.

To illustrate the effect of infrequent trading, consider Figure 2. Panel A plots prices of an asset that starts at $1. Each circle denotes an observation at the end of each quarter. I produced the graphs in Figure 2 by simulation, and deliberately chose one sample path where the prices have gone up and then down to mirror what happened to equities during the 2000s Lost Decade. The

[9] See Ang, Rhodes-Kropf and Zhao (2008).

prices in Panel A appear to be drawn from a series that does not appear excessively volatile; the standard deviation of quarterly returns computed using the prices in Panel A is 0.23.

[Figure 2 here]

The true daily returns are plotted in Panel B of Figure 2. These are much more volatile than the ones in Panel B. Prices go below 0.7 and above 3.0 in Panel B with daily sampling, whereas the range of returns in Panel A is between 1.0 and 2.5 with quarterly sampling. The volatility of quarterly returns, computed from (overlapping) daily data in Panel B is 0.28, which is higher than the volatility of quarterly-sampled returns of 0.23 in Panel A.

For a full comparison, Panel C plots both the quarterly- and daily-sampled returns and just overlays Panel A and Panel B in one picture. Infrequent sampling has caused the estimated volatility estimate using the quarterly-sampled returns to be too low. The same effect also happens to betas and correlations – risk estimates are biased downwards in the presence of infrequent sampling.[10]

### 3.3 Unsmoothing Returns

To account for the infrequent trading bias, we need to go from Panel A of Figure 2, which samples quarterly, to Figure B, which samples daily. That is, the quarterly-observed returns are too smooth and we need to tease out the true, noisier returns. This process is called "*unsmoothing*" or "*desmoothing*," and the first algorithms to do this were developed by David Geltner (1991), a noted professor of real estate at MIT, and Stephen Ross and Randall Zisler (1991). Ross is the same professor who developed multifactor models (see Chapter XX) and Zisler is a real estate professional who started his career as an academic. Ross and Zisler's work

[10] See Geltner (1993) and Graff and Young (1996) for infrequent observation bias on the effect of betas and correlations, respectively. Geltner estimates that betas are understated by a factor of a half for real estate returns.

originally grew out of a series of reports written for Goldman Sachs in the late 1980s. This methodology has been extended in now a very large literature.

Unsmoothing is a *filtering problem*. Filtering algorithms are normally used to separate signals from noise. When we're driving on a freeway and talking on a cell phone, our phone call encounters interference – from highway overpasses and tall buildings – or the reception becomes patchy when we pass through an area without enough cell phone towers. Telecommunication engineers have clever algorithms to enhance the signal, which carries our voice, against all the static. The full transmission contains both the signal and noise, and so the true signal is less volatile than the full transmission. Thus, standard filtering problems are designed to remove noise. The key difference with standard filtering problems is that unsmoothing *adds* noise back to the reported returns to uncover the true returns.

To illustrate the Geltner-Ross-Zisler unsmoothing process, denote the true return at the end of period $t$ as $r_t$, which is unobservable, and the reported return as $r_t^*$, which is observable. Suppose the observable returns follow

$$r_t^* = c + \phi r_{t-1}^* + \varepsilon_t, \tag{1.1}$$

where $\phi$ is the autocorrelation parameter and is less than one in absolute value. Equation (1) is an AR(1) process, where "AR" stands for autocorrelation and the "1" denotes that it captures autocorrelation effects for one lag. Assuming the observed returns are functions of current and lagged true returns (this is called a "*transfer function*" or an "*observation equation*" in the parlance of engineers), we can use equation (1) to invert out the true returns. If the smoothing process only involves averaging returns for this period and the past period, then we can filter the observed returns to estimate the true returns, $r_t$, from observed returns, $r_t^*$, using:

$$r_t = \frac{1}{1-\phi} r_t^* - \frac{\phi}{1-\phi} r_{t-1}^*. \tag{1.2}$$

Equation (2) unsmooths the observed returns. If our assumption on the transfer function is right, the observed returns implied by equation (2) should have zero autocorrelation. Thus, the filter takes an autocorrelated series of observed returns and produces true returns that are close to I.I.D. Note that the variance of the true returns is higher than the observed returns:

$$\text{var}(r_t) = \frac{1+\phi^2}{1-\phi^2} \text{var}(r_t^*) \geq \text{var}(r_t^*), \tag{1.3}$$

since $|\phi| < 1$, so we are adding variance to the observed returns to produce estimates of the true returns.

The unsmoothing process has several important properties:

1. Unsmoothing only affects risk estimates and not expected returns.

   Intuitively, estimates of the mean require only the first and last price observation (with dividends take "total prices" which count reinvested dividends).[11] In Figure 2, we can see that the first and last observations are unchanged by infrequent sampling; thus unsmoothing only changes the volatility estimates.

2. Unsmoothing has no effect if the observed returns are uncorrelated.

[11] Technically taking means of both the right- and left-sides in equation (2) results in the same means in large samples.

In many cases, reported illiquid asset returns are autocorrelated because illiquid asset values are appraised. The appraisal process induces smoothing because appraisers use, as they should, both the most recent and comparable sales (which are transactions) together with past appraised values (which are estimated, or perceived, values). The artificial smoothness from the appraisal process has pushed many in real estate to develop pure transactions-based, rather than appraisal-based indexes.[12] Autocorrelation also results from more shady aspects of subjective valuation procedures – the reluctance of managers to mark to market in down markets.

In many cases, we expect the true illiquid asset returns to be autocorrelated as well.[13] Illiquid asset markets – like real estate, private equity, timber plantations, and infrastructure – are markets where information is not available to all participants, information does not spread rapidly, and capital cannot be immediately deployed into new investments. Informationally inefficient markets with slow-moving capital are characterized by persistent returns.[14]

3. Unsmoothing is an art.

   The unsmoothing example in equations (1)-(2) uses the simplest possible autocorrelated process, an AR(1), to describe reported returns. Many illiquid assets have more than first-order lag effects. Real estate, for example, has a well-known fourth-order lag arising from

[12] This literature includes both repeat-sales methodologies (see Goetzmann (1992)) and constructing indexes using only transactions (see Gatzlaff and Geltner (1998) and Fisher, Geltner and Pollakowski (2007)). The aggregation process in constructing indexes of illiquid asset returns induces further smoothing. Indexes combine many individual indications of value, either market transactions or appraised values, and typically the values are appraised at different points throughout the year. Note that if $\phi = 0$, then equations (2) and (3) coincide and unsmoothed returns are exactly the same as reported returns. Figure 2, which shows the effects of infrequent observations, is produced with a year-on-year autocorrelation of 0.4.

[13] When the true returns are autocorrelated, the horizon matters in stating volatilities, correlations, and Sharpe ratios. From point 1, the means are unaffected. See Lo (2002) for formulas to appropriately convert the risk measures for different horizons.

[14] See Duffie (2010).

many properties being reappraised only annually.[15] A good unsmoothing procedure takes a time-series model that fits the reported return data well and then with a general transfer function assumption, the filter for true returns in equation (2) becomes a very complicated function of present and past lagged observed returns.[16] Doing this properly requires good statistical skills. It also requires underlying economic knowledge of the structure of the illiquid market to interpret what is a reasonable lag structure, and to judge how much unsmoothing is required.

### *Unsmoothed Real Estate Returns*

To illustrate the effects of unsmoothing,[17] Figure 3 plots direct real estate returns from the National Council of Real Estate Investment Fiduciaries (NCREIF), which constructs an institutional property index from data reported by its members. Because this is an appraisal index, NCREIF real estate returns are highly autocorrelated. From March, 1978 to December, 2011, the first-order autocorrelation of NCREIF returns is 0.78. The raw reported data is shown in the solid line. I graph unsmoothed returns in the squares applying the filter of equations (1) and (2). All returns are at the quarterly frequency.

[Figure 3 here]

Unsmoothing produces a dramatic effect. The minimum reported return during the real estate downturn in the early 1990s is -5.3% during the quarter ending December 1991. The corresponding unsmoothed return is -22.6%. During the financial crisis, NCREIF returns reached

[15] This is noted in the seminal Geltner (1991) and Ross and Zisler (1991) papers.

[16] We want an ARMA(p,q) model, which captures the effect of *p* lagged autocorrelated terms (the "AR" effect for *p* lags) and where innovations to those returns in past periods continue to have an effect on present returns. The latter are referred to as moving average terms (the "MA" effects for *q* lags). Both Geltner (1991) and Ross and Zisler (1991) consider richer time-series processes than just an AR(1). Okunev and White (2003) develop unsmoothing algorithms to hedge fund returns with higher-order autocorrelation corrections.

[17] Unsmoothing corrections produce similar effects in other illiquid markets. Campbell (2008), for example, estimates that unsmoothing increases the volatility of art market returns from 6.5% to 11.5%.

a low of -8.3% in December 2008. The unsmoothed return during this quarter is -36.3%. The volatility of the raw NCREIF returns is 2.25% per quarter, whereas the volatility of the unsmoothed returns is 6.26% per quarter. This approximates the volatility of stock returns, which is around 7.5% per quarter. Correlation (and hence beta) estimates are also affected by unsmoothing: the correlation of raw NCREIF returns with the S&P 500 is 9.2% and this rises to 15.8% once the unsmoothing correction is applied.

### 3.4 Selection Bias

The tendency of returns only to be observed when underlying asset values are high is called "*sample selection bias*."

Buildings tend to be sold when their values are high – otherwise, many sellers postpone sales until property values recover. This causes more transactions to be observed when the underlying real estate values are high.

In private equity, selection bias is acute. In buyout funds, companies are taken public only when stock values are high. Many venture capital investments are structured over multiple rounds. Better-performing companies tend to raise more money in more rounds. The venture capitalist tends to sell a small company, and the transaction is recorded, when the company's value is high. Distressed companies are usually not formally liquidated and these "zombie" companies are often left as shell companies. When observing old companies without recent transactions, it is not clear whether these companies are alive and well or whether they are zombies.

To illustrate the selection bias problem, consider Figure 4 which is adapted from Korteweg and Sorensen (2010). Panel A shows the full universe of returns of an illiquid asset marked by dots. These returns (on the *y*-axis) are plotted contemporaneous with market returns (on the *x*-axis). In the full universe, there is no alpha and the intercept of the line summarizing the relationship between the illiquid asset and the market goes through the origin (this line is called the Security Market Line (SML), see Chapter XX). The slope of the SML is the beta of the illiquid asset and is a measure of risk.

[Figure 4 here]

Panel B illustrates the sample selection problem. Bad returns, which are shaded gray, are not observed in the databases – we only record transactions when prices are high. Now only the black dots are reported. An estimated SML fitted to these observed returns yields a positive alpha when the true alpha is zero. The slope of the fitted SML is flatter than the slope of the true SML in panel A, and hence we underestimate beta. When we compute the volatility of the observed returns, we only count those returns that are high, and so the volatility estimate is biased downwards. Thus, we over-estimate expected return, and we under-estimate risk as measured by beta and volatility.

The statistical methodology for addressing selection bias was developed by James Heckman (1979), who won the Nobel Prize in 2000 for inventing these and other econometric techniques. Studies which use models to correct for these biases do not take such an extreme view as Figure 4: they allow the threshold above which returns are observed to vary over time and depend on company or property-level characteristics.[18] The model of risk is sometimes extended to multifactor models (see Chapter XX), rather than just using the market portfolio as the sole risk factor.

[18] See also Cochrane (2005) for selection bias models applied to venture capital and Fisher et al. (2003) for real estate.

The effect of selection bias can be enormous. Cochrane (2005) estimates an alpha for venture capital log returns of over 90% not taking into account selection bias, which reduces to -7% correcting for the bias. Korteweg and Sorensen (2005) estimate that expected returns for the same asset class are reduced downwards by 2-5% per month (arithmetic returns) taking into account selection bias. The effect of selection bias in real estate is smaller, perhaps because the underlying volatility of real estate returns is lower than private equity. Fisher et al. (2003) implement selection bias corrections for real estate. They estimate that average real estate returns reduce from 1.7% to 0.3% and standard deviation estimates increase by a factor of 1.5. The small means of real estate returns are due to their sample period of 1984-2001, which includes the real estate downturn in the early 1990s and in the early 2000s. They miss the bull market in real estate during the mid-2000s.

### 3.5 Summary

Treat reported illiquid asset returns very carefully. Survivors having above-average returns, infrequent observations, and the tendency of illiquid asset returns to be reported only when underlying valuations are high will produce return estimates that are overly optimistic and risk estimates that are biased downwards. Put simply, reported returns of illiquid assets are too good to be true.

## 4 Illiquidity Risk Premiums

Illiquidity risk premiums compensate investors for the inability to access capital immediately. They also compensate investors for the withdrawal of liquidity during illiquidity crises.

### *Harvesting Illiquidity Risk Premiums*

There are four ways an asset owner can capture illiquidity premiums:

1. By setting a *passive allocation* to *illiquid asset classes*, like real estate;
2. By choosing securities within an asset class that are more illiquid, that is by engaging in *liquidity security selection;*
3. By acting as a *market maker* at the individual security level;
4. By engaging in *dynamic strategies* at the aggregate portfolio level.

Economic theory states that there should be a premium for bearing illiquidity.[19] Theory also states that the premium can be large or small. It is perhaps surprising that the premium for illiquidity could be small.[20] In models where illiquidity risk has small or no effect on prices, illiquidity washes out across individuals. A particular individual may be affected by illiquidity – illiquidity can crimp his consumption, or affect his asset holdings (as in the asset allocation model with illiquidity risk I present below) – but other agents will not be constrained, or they trade at different times. Different agents share risk among themselves, which mutes the impact of illiquidity. Thus, in equilibrium the effects of illiquidity can be negligible.[21]

Whether the illiquidity risk premium is large or small is an empirical question.

## 4.1 Illiquidity Risk Premiums Across Asset Classes

Figure 5 is from Antti Ilmanen's (2011) wonderful book, *Expected Returns*, and plots average returns on illiquidity estimates. The average returns are computed from (reported) data over 1990 to 2009. The illiquidity estimates represent Ilmanen's opinions. Some private equity investments are more liquid than certain hedge funds, and some infrastructure investments are much less

[19] This large literature begins with a seminal contribution by Demsetz (1968). See summary articles by Hasbrouck (2007) and Vayanos and Wang (2012).
[20] Lo, Mamaysky and Wang (2004) and Longstaff (2009), among others, argue that the illiquidity premium should be large.
[21] For models of this kind, see Constantinides (1986), Vayanos (1998), Garlenau (2009), and Buss, Uppal and Vilkov (2012).

liquid than private equity, so it is hard to pigeon-hole these asset classes in terms of illiquidity. Nevertheless, Figure 5 seems to suggest a positive relation between how illiquid an asset class is and its expected return. Figure 5 represents "conventional" views among most market participants that there is a reward to bearing illiquidity across asset classes.

This conventional view is flawed for the following reasons:

1. Illiquidity biases

   As Section 3 shows, reported data of illiquid assets cannot be trusted. The various illiquidity biases – survivorship bias, sampling at infrequent intervals, and selection bias – result in the expected returns of illiquid asset classes being overstated using raw data.

2. Ignores risk

   Illiquid asset classes contain far more than just illiquidity risk. Adjusting for these risks makes illiquid asset classes far less compelling. Chapter XX showed that the NCREIF real estate index (despite its raw returns being overly optimistic due to illiquidity biases) is beaten by a standard 60% equity and 40% bond portfolio. Chapters XX and XX will show that the average hedge fund and private equity fund, respectively, provide zero expected excess returns. In particular, after adjusting for risk, most investors are better off investing in the S&P 500 than in a portfolio of private equity funds.

3. There is no "market index" for illiquid asset classes

   Indexes of illiquid asset classes are not investable. An asset owner *never* receives the NCREIF return on a real estate portfolio. The same is true for most hedge fund indexes and private equity indexes. In liquid public markets, large investors can receive index market returns and pay close to zero in fees. In contrast, NCREIF is not investable as it is impossible to buy all the underlying properties in that index. Since all asset owners own considerably fewer properties than the thousands included in NCREIF, they face considerably more idiosyncratic risk. While this large amount of idiosyncratic risk can boost returns in some cases, it can also lead to the opposite result. Returns to illiquid asset investing can be far below a reported index.

4. You cannot separate factor risk from manager skill

   Tradeable and cheap index funds in bond and stock markets allow investors to separate systematic returns (or factor returns, see Chapter XX) from management talent. In illiquid markets, no such separation is possible: *investing in illiquid markets is always a bet on management talent*. The agency issues in illiquid asset markets are first-order problems, and I discuss them in Part III of this book. Agency issues can, and often do, overwhelm any advantages that an illiquidity risk premium may bring.

Taking into account data biases, the evidence for higher average returns as asset classes become more illiquid is decidedly mixed, as Ang, Goetzmann and Schaefer (2011) detail.[22] But while there do not seem to be significant illiquidity risk premiums *across* classes, there are large illiquidity risk premiums *within* asset classes.

## 4.3 Illiquidity Risk Premiums Within Asset Classes

Within all the major asset classes, securities that are more illiquid have higher returns, on average, than their more liquid brethren.

### *U.S. Treasuries*

A well-known liquidity phenomenon in the U.S. Treasury market is the *on-the-run/off-the-run bond spread*. Newly auctioned Treasuries (which are "on the run") are more liquid and have

[22] Nevertheless, there are common components in illiquidity conditions across asset classes: when U.S. Treasury bond markets are illiquid, for example, many hedge funds tend to do poorly. See, for example, Hu, Pan and Wang (2012).

higher prices, and hence lower yields, than seasoned Treasuries (which are "off the run").[23] The spread between these two types of bonds varies over time reflecting time-varying liquidity conditions in Treasury markets.[24] (For more details see Chapter XX.)

There were pronounced illiquidity effects in Treasuries during the 2008-2009 financial crisis. Treasury bonds and notes are both issued by the U.S. Treasury and are identical except that bonds have maturities of 20-30 years and notes originally have maturities of 1-10 years. But after 10 years, a Treasury bond originally carrying a 20 year maturity is the same as a Treasury note. If the maturities are the same, whether the U.S. Treasury is a bond or a note should make no difference. During the financial crisis Treasury bonds traded lower than Treasury notes by more than 5% – all illiquidity effects on otherwise identical securities. In one of the world's most important and liquid markets, these are very large illiquidity effects.

### *Corporate Bonds*

Corporate bonds that trade less frequently or have larger bid-ask spreads have lower returns. Chen, Lesmond and Wei (2007) find that illiquidity risk explains 7% of the variation across yields of investment-grade bonds. Illiquidity accounts for 22% of the differences in yields in junk bonds and for these bonds, a one basis point rise in bid-ask spreads increases yield spreads by more than two basis points.[25]

[23] The on-the-run bonds are more expensive because they can be used as collateral for borrowing funds in the repo market. This is called "*specialness*." See Duffie (1996).
[24] See Goyenko, Subrahmanyam and Ukhov (2011)
[25] See also Chapter XX, Bao Pan and Wang (2011), Lin, Wang and Wu (2011), and Dick-Nielsen, Feldhutter and Lando (2012).

### *Equities*

A large literature finds that many illiquidity variables predict returns in equity markets, with less liquid stocks having higher returns.[26] These variables include bid-ask spreads, volume, volume signed by whether trades are buyer or seller initiated, turnover, the ratio of absolute returns to dollar volume (commonly called the "Amihud measure" based on his paper of 2002), the price impact of large trades, informed trading measures (which gauge adverse, informed trading, see below), quote size and depth, the frequency of trades, how often there are "zero" returns (in more liquid markets returns will bounce up and down), and return autocorrelations (which are a measure of stale prices).[27] These are all illiquidity characteristics, which are properties unique to an individual stock. There are also illiquidity risk betas. These are covariances of stock returns with illiquidity measures, like market illiquidity or signed volume.

Estimates of illiquidity risk premiums in the literature range between 1-8% depending on which measure of illiquidity is used. However, Ben-Rephael, Kadan and Wohl (2008) report that these equity illiquidity premiums have diminished considerably – for some illiquidity measures the risk premiums are now zero! In pinksheet stock markets, which are OTC equity markets, Ang, Shtauber and Tetlock (2011) find an illiquidity risk premium of almost 20% compared to comparable listed equity illiquidity risk premiums of approximately 1%.

### *Illiquid Assets*

There are higher returns to hedge funds that are more illiquid, in the sense that they place more restrictions on the withdrawal of capital (called lockups, see Chapter XX) or for hedge funds

[26] See the summary article by Amihud, Mendelson and Pedersen (2005).
[27] Some of these illiquidity measures produce spreads in expected returns opposite to an illiquidity risk premium, Stocks with higher than average (normalized) volume, for example, tend to have lower future returns as shown by Gervais, Kaniel and Mingelgrin (2001).

whose returns decrease when liquidity dries up.[28] Franzoni, Nowak and Phalippou (2012) show that there are significant illiquidity premiums in private equity funds, with illiquidity effects of 3% for the typical private equity fund (for further details, see Chapter XX). In real estate, Liu and Qian (2012) construct illiquidity measures of price impact and search costs for U.S. office buildings. They find a 10% increase in these illiquidity measures leads to a 4% increase in expected returns for these properties.

### *Why Illiquidity Risk Premiums Manifest Within but not Across Asset Classes*

To my knowledge, we have yet to develop formal equilibrium models explaining the large illiquidity risk premiums within asset classes, but not across asset classes.

Perhaps the reason is limited integration across asset classes. There are significant impediments to switching capital and investment strategies seamlessly even across liquid stock and bond markets.[29] Investors put asset classes into different silos, and rarely treat them consistently as a whole. This happens on both the sell-side, where fixed income, equity desks, and other divisions rarely talk with each other, and on the buy-side where each asset class is managed by separate divisions. (Canada Pension Plan's factor investing strategy is a notable exception to this, as I discuss in Chapter XX.) The potential mispricing of illiquidity across asset classes could reflect institutional constraints, slow-moving capital, and limits to arbitrage.[30]

On the other hand, perhaps asset class illiquidity risk premiums are small because investors overpay for illiquid asset classes; they chase the illusions of higher returns, and bid up the prices of these illiquid assets causing the illiquidity premiums to go away. In contrast, within asset

[28] See Aragon (2007) and Sadka (2010), respectively.
[29] See Kapadia and Pu (2012) for evidence of lack of integration across stock and bond markets.
[30] See Merton (1987), Duffie (2010), and Shleifer and Vishny (1997), respectively.

classes – especially the more liquid stock and bond markets – highly illiquidity risk-averse investors are willing to pay for the privilege to trade as soon as they desire.

## 4.4 Market Making

A market maker supplies liquidity by acting as an intermediary between buyers and sellers.[31] Liquidity provision is costly – market makers need capital to withstand an onslaught of many repeated buys or sells. They could be transacting with investors who have superior information. Thus, market makers need to be compensated for supplying immediacy and demand price concessions: they buy at low prices and sell at prices around "fair value." Investors transacting with the market maker pay the *bid-ask spread*.

In liquid stock and bond markets, market making is now synonymous with high frequency trading – traders who build massive computer infrastructure to submit buy and sell orders within fractions of a second. Recent estimates put the share of dollar trading volume on U.S. equity exchanges due to high frequency traders at above 70%.[32] There are many successful hedge funds whose specialty is high frequency trading (see Chapter XX).

Many asset owners cannot collect illiquidity risk premiums by building high frequency trading systems, nor would they wish to enter this business (directly or indirectly). But there is a way large asset owners can do a low-frequency version of market making.

Dimensional Funds Advisors (DFA) is a funds management company that started in 1981 by specializing in small-cap equities. This company's strategies have deep roots in academic factor models, and its founders David Booth and Rex Sinquefield, roped in the big guns of the finance literature, Fama, French, and others, in building the company. From the start, DFA positioned

[31] O'Hara (1995) provides a summary of theoretical models of market making.
[32] See Zhang (2010).

itself as a liquidity provider, especially of small stocks, and market making was an integral part of its investment strategy. [33] When other investors seek to offload large amounts of small stocks, DFA takes the other side and buys at a discount. Similarly, DFA offers large blocks of small-cap equities sells at a premium to investors who demand immediate liquidity.

Large asset owners, like sovereign wealth funds and large pension funds, are in a position to act as liquidity providers, especially in more illiquid markets. They can accept large blocks of bonds, shares, or even portfolios of property at discount and sell these large blocks at premiums. They can do this by calculating limits within their (benchmark tracking error) constraints of how much they are willing to transact. That is, they can provide liquidity in different securities up to a certain amount so that they do not stray too far from their benchmarks. Buyers and sellers will come to them as they develop reputations for providing liquidity.

### *Secondary Markets for Private Equity and Hedge Funds*

Exchanges for secondary transactions in hedge funds and private equity have sprung up, but these markets are still very thin. Many transactions also do not take place on organized secondary market platforms.

There are two forms of secondary markets in private equity. First, in secondary (and tertiary) market buyout markets, private equity firms trade private companies with each other. These markets have blossomed: in 2005, secondary buyouts represented around 15% of all private equity buy-out deals.[34] From the perspective of asset owners (limited partners, or LPs), this market provides no exit opportunities from the underlying private equity funds and is at worst a merry-go-round of private equity firms swapping companies in circular fashion. At best, more

[33] See Keim (1999) and the Harvard Business School case study, Dimensional Fund Advisers, 2002, written by Randolph Cohen.
[34] Report of the Committee on Capital Markets Regulation, 2006,

transactions at market prices (assuming there is no finagling between the transacting funds) allow asset owners to better value their illiquid investments. The limited partners receive some cash earlier when a company in their fund's portfolio is sold to another private equity firm, even though they are still stuck in the private equity fund.

Secondary markets for limited partners are much smaller and more opaque. Even industry participants acknowledge this market "still remains relatively immature … and still represents a very small percentage of the primary market."[35] Bid-ask spreads in these transactions are enormous. As Cannon (2007) notes, the secondary market for LPs was dominated in the 1990s by distressed sellers. Specialized firms on the other side received discounts of 30-50%; there was a reason these firms were called "vultures." In the 2000s, discounts fell to below 20%, but shot up during the financial crisis. Harvard University found this out when it tried to disinvest in private equity funds during 2008 and faced discounts of 50%.

Discounts for hedge funds are much smaller than private equity. This reflects the fact that hedge funds investors can, in most cases, access capital at pre-determined dates after lockups have expired and notice periods have been given (unless the hedge fund imposes gates). Reflecting this underlying greater liquidity, hedge fund discounts in secondary markets in 2007 and 2008 were around 6-8%.[36] (A few hedged funds that are closed to new investors actually trade at premiums.)

The nascent secondary markets for private equity and hedge funds represent tremendous opportunities for large asset owners to supply liquidity. Secondary private equity is like second-hand cars that are still brand new. When you drive a new car off the lot, it immediately

[35] From the introduction to Luytens (2008) written by Andrew Sealey and Campbell Lutyens. Is Campbell Lutyens one person? I thought it was a firm.
[36] See Ramadorai (2012).

depreciates by a quarter, even though it is exactly the same as a car sitting in the dealer's inventory. Secondary private equity is still private equity, and you can get it a lot cheaper than direct from the dealer.

### *Adverse Selection*

A market maker faces a risk that a buyer has non-public information and the stock that is selling could be too low. Or, a buyer could continue to buy and further increase the price. In this case, the market maker has sold too early and too low. This is *adverse selection*. Glosten and Milgrom (1985) and Kyle (1985) – the papers that started the market-making microstructure literature – developed theories of how the bid-ask spread should be set to incorporate the effects of adverse selection. DFA provides some examples of how to counter adverse selection. To avoid being exploited, DFA trades with counterparties that fully disclose their information on stocks. At the same time, DFA itself operates in a trustworthy way by not front running and by not manipulating prices.

## 4.5 Rebalancing

The last way an asset owner can supply liquidity is through dynamic portfolio strategies. This has a far larger impact on the asset owner's total portfolio than liquidity security selection or market making because it is a top-down asset allocation decision (see Chapter XX for factor attribution).

*Rebalancing is the simplest way to provide liquidity*, as well as being the foundation of all long-horizon strategies (see Chapter XX). Rebalancing forces asset owners to buys assets at low prices when others want to sell. Conversely, rebalancing automatically sheds assets at high prices, giving them to investors who want to buy at elevated levels. Since rebalancing is countercyclical, it supplies liquidity. Dynamic portfolio rules, especially those anchored by simple valuation rules (see Chapters XX and XX) extend this further – as long as they buy when others want to sell and vice versa. It is especially important to rebalance illiquid asset holdings too, when given the chance (see also below).

Purists will argue that rebalancing is not strictly liquidity provision; rebalancing is an asset management strategy. Rebalancing, in fact, requires liquid markets to implement. But prices exhibit large declines often because of blowouts in asymmetric information, or because funding costs rapidly increase so that many investors are forced to offload securities – some of the key reasons giving rise to illiquidity listed at the start of Section 2. Brunnermeier (2009) argues that these effects played key roles in the meltdown during the financial crisis. In the opposite case, rebalancing makes available risky assets to new investors, potentially with lower risk aversions than existing clientele or those who chase past high returns, or to investors who load up on risky assets when prices are high because they have abundant access to leverage and they perceive asymmetric information is low. In this general framework, *rebalancing provides liquidity*.

Large asset owners give up illiquidity premiums by sheepishly tracking standard indexes. When indexes change their constituents, asset owners demand liquidity as they are forced to follow these changes. Index inclusion and exclusion induce price effects of 3-5%, and these effects have become stronger in more recent data.[37] Large asset owners should be collecting index reconstitution premiums instead of paying them. They can do this by using their own proprietary benchmarks. Candidate indexes could emphasize illiquidity security characteristics, but more generally would be built around harvesting factor risk premiums (see Chapter XX). Even an

[37] See Chapter XX and the literature on index reconstitution effects summarized by Ang, Goetzmann and Schaeffer (2011).

index without illiquidity tilts allows asset owners to harvest a liquidity premium collected from all the other investors forced to track standard indexes.

### 4.6 Summary

Of all the four ways to collect an illiquidity premium summarized at the beginning of this section: (1) holding passive allocations to illiquid asset classes; (2) holding less liquid securities within asset classes; (3) market making at the individual security level; and (4) dynamic rebalancing at the aggregate level, the last of these is simplest to implement and has the greatest impact on portfolio returns.

## 5 Portfolio Choice with Illiquid Assets

Given that illiquid assets do not have tradeable index returns, investors face agency issues and need skill to evaluate and monitor managers, and need sufficient resources to handle investments over long periods, it may be appropriate to use individual-specific illiquidity premiums. Computing these requires asset allocation models with liquid and illiquid assets. These models also prescribe an optimal amount of illiquid assets to hold.

Practitioners generally use one-period investment models – usually the restrictive Markowitz (1952) mean-variance model with ad-hoc adjustments (yes, most of the industry is still using models from the 1950s, see Chapter XX) – which are inappropriate for illiquid asset investing. The fact that you cannot trade an illiquid asset now, but will do so in the future makes illiquid asset investing a dynamic, long-horizon problem. There are two important aspects of illiquidity –

large transaction costs and long times between trading – that have been captured in portfolio choice models with illiquid assets.[38]

## 5.1 Asset Allocation with Transactions Costs

George Constantinides (1986) was the first to develop an asset allocation model where the investor had to pay transactions costs to trade. Selling $100 of equities, for example, results in a final position of $90 with 10% transactions costs. Not surprisingly, the investor trades infrequently – to save on transactions costs. Constantinides proved that the optimal strategy is to trade whenever risky asset positions hit upper or lower bounds. Within these bounds is an interval of no trading. The no-trading band straddles the optimal asset allocation from a model that assumes you can continuously trade without frictions (the Merton (1971) model).[39]

The no-trade interval is a function of the size of the transactions costs and the volatility of the risky asset. Constantinides estimates that even for transactions costs of 10%, there are no-trade intervals greater than 25% around an optimal holding of approximately 25% for a risky asset with a 35% volatility. (I bet Harvard wished it could have received just a 10% discount when it tried to sell its private equity investments.) That is, the asset owner would not trade between [0%, 50%] – indeed, very large fluctuations in the illiquid asset position. Illiquid asset investors should expect to rebalance very infrequently.

Constantinides' model can be used to compute an illiquidity risk premium, defined as the expected return of an illiquid asset required to bring the investor the same level of utility as in a frictionless setting. This is the risk premium the investor demands to bear the transactions costs and is a certainty equivalent calculation (see Chapter XX). For transactions costs of 15% or

[38] Parts of this are based on Ang (2011) and Ang and Sorensen (2012).
[39] Chapter XX discusses extensions of Constantinides (1986) to double bands, contingent bands, and rebalancing to the edge or center of the bands.

more, the required risk premium exceeds 5%. Compare this value with (the close to) zero additional excess returns, on average, of the illiquid asset classes in data.

A major shortcoming of the transaction costs models is that they assume trade is always possible by paying a cost. This is not true for private equity, real estate, timber, or infrastructure. Over a short horizon, there may be no opportunity to find a buyer. Even if a buyer can be found, you need to wait for due diligence and legal transfer to be completed.[40] Many liquid assets also experienced liquidity freezes during the financial crisis where no trading – at any price – was possible because no counterparties could be found.

## 5.2 Asset Allocation with Infrequent Trading

In Ang, Papanikolaou and Westerfield (2012), I develop an asset allocation model the investor can transact illiquid assets only at the arrival of randomly occurring liquidity events. This notion of illiquidity is that usually illiquid assets are just that – illiquid and cannot be traded. But when the liquidity event comes, investors can trade the illiquid assets.

I model the arrival of liquidity events by a Poisson arrival process with intensity $\lambda$. The interval between liquidity events is $1/\lambda$. For real estate or private equity, intervals between trading would occur every 10 years or so, so $\lambda = 1/10$. As $\lambda$ increases to infinity, the opportunities to rebalance become more and more frequent and in the limit approach the standard Merton (1981) model where trading occurs continuously. Thus, $\lambda$ indexes a range of different illiquidity outcomes.

[40] For some illiquid assets, investors may not be even willing to transact immediately for one cent; some investments do not have liability limited at zero. For example, at June 30, 2008, a real estate investment by CalPERS was valued at negative $300 million! See Corkery, M., C. Karmin, R. L. Rundle, and J. S. Lublin, "Risky, Ill-Timed Land Deals Hit CalPERS," Wall Street Journal, Dec. 17, 2008.

Poisson arrival events have been used to model search-based frictions since Peter Diamond (1982). Diamond was awarded the 2010 Nobel Prize. In 2011, he was nominated to serve on the Federal Reserve Board of Governors, but Republican opposition prevented his confirmation.

Illiquidity risk causes the investor to behave in a more risk-averse fashion toward both liquid and illiquid assets. Illiquidity risk induces time-varying, endogenous risk aversion. Harvard discovered in 2008 that although it is wealthy, it cannot "eat" illiquid assets. Illiquid wealth and liquid wealth are not the same; agents can only consume liquid wealth. Thus, the solvency ratio of illiquid to liquid wealth affects investors' portfolio decisions and payout rules – it becomes a state variable that drives investors' effective risk aversion.

The takeaways from the Ang, Papanikolaou and Westerfield (2012) model are:

### *Illiquidity Markedly Reduces Optimal Holdings*

Start with the bottom line in Panel A of Table 6, which reports a baseline calibration where the investor holds 59% in a risky asset that can always be traded. This weight is very close to the standard 60% equity allocation held by many institutions. As we go up the rows, the asset becomes more illiquid. If the risky asset can be traded on average every six months, which is the second last line, the optimal holding of the illiquid asset contingent on the arrival of the liquidity event is 44%. When the average interval between trades is five years, the optimal allocation is 11%. For ten years, this reduces to 5%. Illiquidity risk has a very large effect on portfolio choice.

[Table 6 here]

### *Rebalance Illiquid Assets to Positions Below the Long-Run Average Holding*

In the presence of infrequent trading, illiquid asset wealth can vary substantially and is right-skewed. Suppose the optimal holding of illiquid assets is 0.2 when the liquidity event arrives.

The investor could easily expect illiquid holdings to vary from 0.1 to 0.35, say, during non-rebalancing periods. Because of the right-skew, the average holding of the illiquid asset is 0.25, say, and is greater than the optimal rebalanced holding. The optimal trading point of illiquid assets is lower than the long-run average holding.

### *Consume Less with Illiquid Assets*

Payouts, or consumption rates, are lower in the presence of illiquid assets than when only comparable liquid assets are held by the investor. The investor cannot offset the risk of illiquid assets declining when these assets cannot be traded. This is an *unhedgeable* source of risk. The investor offsets that risk by eating less.

### *There are No Illiquidity "Arbitrages"*

In a mean-variance model, two assets with different Sharpe ratios and perfect correlations produce positions of plus or minus infinity. This is a well-known bane of mean-variance models, and professionals employ lots of ad-hoc fixes, usually (sometimes arbitrary) constraints, to prevent this from happening. This does not happen when one asset is illiquid – there is no arbitrage. Investors do not load up on illiquid assets because these assets have illiquidity risk and cannot be continuously traded to construct an "arbitrage."

### *Investors Must Demand High Illiquidity Hurdle Rates*

How much does an investor need to be compensated for illiquidity? In Panel B of Table 6, I compute premiums on an illiquid asset required by an investor to bear illiquidity risk. I define the illiquidity premium as a certainty equivalent. Suppose an investor holds two liquid assets and replaces one liquid asset with an illiquid asset that is identical to the liquid one, except it cannot be immediately traded. The illiquidity premium is the increase in the expected return of the illiquid asset so that the investor has the same utility as the case when all assets are liquid.

When liquidity events arrive every six months, on average, an investor should demand an extra 70 basis points. (Some hedge funds have lockups around this horizon.) When the illiquid asset can be traded once a year, on average, the illiquidity premium is approximately 1%. When you need to wait ten years, on average, to exit an investment, you should demand a 6% illiquidity premium. That is, investors should insist that private equity funds generate returns 6% greater than public markets to compensate for illiquidity. As Section 3 discusses, most illiquid assets are not generating average returns above these hurdle rates (see also Chapter XX).

The Ang, Papankiolaou and Westerfield (2012) model is highly stylized. Given the other issues the model misses, like agency conflicts of interest (see Chapter XX), cash flow management issues of capital calls and distributions, and asset-liability mismatches, the true illiquidity hurdle rates are even higher than those reported in Table 6.

### 5.3 Summary

Portfolio choice models with illiquid assets recommend holding only modest amounts of illiquid assets. Investors should demand high illiquidity risk premiums.

## 6 Liquidating Harvard Redux

### 6.1 The Case for Illiquid Investing

Large, long-term investors often cite their large amounts of capital and their long horizons as rationales for investing in illiquid assets. Size and patience are necessary conditions for illiquid asset investing, but not sufficient ones; these conditions simply aren't adequate justifications in themselves. Since illiquid asset classes do not offer high risk-adjusted returns, the case for passively holding illiquid asset classes is not compelling. Illiquid investing also poses huge agency problems; asset owners, for example, find it tough to monitor external managers. Many institutions face "fiefdom risk" as illiquid assets are run as separate empires within an organization.

In addition, investors in illiquid markets face high idiosyncratic risk because there is no "market" portfolio. It is exactly this large idiosyncratic risk, however, that is the most compelling reason for investing in illiquid assets.

Suppose you are a skilled investor (assume you have true alpha, see Chapter XX) and have a choice between investing in (1) a market where prices quickly reflect new information, almost everyone sees the same information, and news gets spread around very quickly, or (2) a market where information is hard to analyze and even harder to procure, only a select few have good information, and news takes a long time to reach everyone. Obviously you pick (2). This, in a nutshell, is the Swensen (2009) justification for choosing illiquid assets. The argument is not that illiquid asset classes have higher risk-adjusted returns – empirical evidence suggests they don't.

Illiquid asset investing allows an investor to transfer idiosyncratic risk from liquid equity and bond markets, which are largely efficient, to markets where there are large information asymmetries, transactions costs are punishing, and the cross sections of alpha opportunities are extremely disperse. These are the markets, in other words, where you can make a killing!

The Swensen case crucially relies on one word: "skilled." Whereas skilled investors can find, evaluate, and monitor these illiquid investment opportunities, assuming they have the resources to take advantage of them, unskilled investors will be duped. If you are unskilled, you lose. Harvard, Yale, Stanford, MIT, and a few other select endowments have the ability to select superior managers in illiquid markets because of their size, their relationships, and their ability to support these managers through long investment cycles. What about the others? An endowment specialist says, "It's a horror show. [Performance has] been flat to even negative. The strong get stronger and the weak get stuck with non-top quartile managers and mediocre returns and high fees."[41]

## 6.2 Investment Advice for Endowments

Thomas Gilbert and Christopher Hrdlicka at the University of Washington are probably the world's only endowment management theorists. In a 2012 paper, they provocatively argue that the optimal allocation policy for successful universities is to hold large amounts of fixed income, not risky assets, and by extension not illiquid risky assets.

Gilbert and Hrdlicka model universities as creators of "social dividends," which are research and teaching. Universities can invest internally, in research and teaching projects, or they can invest externally through the endowment. If the endowment is taking on risk, through equities, this signals the university does not have enough good internal risky projects generating social dividend, so risk is taken on externally. If the endowment is invested in safe assets, through bonds, the university takes on risk through internal research and teaching projects. Gilbert and Hrdlicka argue that a university endowment's large investment in risky assets is a signal it that it does not have enough fruitful research and teaching assignments!

Harvard, with its large endowment heavily invested in risky illiquid assets, would take issue with Gilbert and Hrdlicka. An endowment allows a university to be more independent, rather than depend entirely on grants from government or private foundations. As Dershowitz argues, the

[41] Quoted by Stewart, J. B., A Hard Landing for University Endowments, New York Times, October 12, 2012.

endowment could be used as a rainy day account to be tapped precisely during times like 2008. Harvard's endowment has historically yielded a predictable stream of cash for operating budgets, but 2008 blew this predictability away. Harvard claims its endowment allows for future generations to share in its riches, saying, "Although their specific uses vary, endowment funds have a common purpose: to support activities not just for one year, or even one generation, but in perpetuity."[42] The price of education, however, has been rising in real terms (see Chapter XX) and if education is costlier in the future than in the present, being stingy on research and teaching now makes no sense because it substitutes a more expensive good in the future for a cheaper one today.[43]

Henry Hansmann, a professor at Yale Law School, describes large private universities as "institutions whose business is to run large pools of investment assets… They run educational institutions on the side, that can expand and contract to act as buffers for investment pools."[44] He contends that a large part of why universities like large endowments is prestige, pursued as its own objective. Journalist Kevin Carey puts it another way, echoing the cadences of the *Book of Common Prayer* when he says that large endowments per se are "aspiration without limit, accumulation without end."[45]

## 6.3 Liquidate Harvard?

Did Harvard generate excess returns, or an illiquidity risk premium, from its large investments in illiquid, alternative assets? Yes. Harvard could extract value from illiquid asset investing not

---

[42] "About HSPH: Endowment Funds: What Are Endowment Funds?" Harvard School of Public Health, http://www.hsph.harvard.edu/about/what-are-endowment-funds
[43] See Hansmann (1990).
[44] "Q&A. Modest Proposal. An Economist Asks, Does Harvard Really Need $15 Billion?" New York Times, Aug. 2, 1998.
[45] Kevin Carey, "The 'Veritas' About Harvard," Chronicle of Higher Education, Sept. 28, 2009.

because illiquid asset classes have a large risk premium, but because it is a skilful investor. And it is one of the few investors able to do so.

But this doesn't help Harvard solve its cash crunch. The worst failing of Harvard was in basic asset-liability management. Even without using the asset allocation models with illiquidity risk or the advice given by Gilbert and Hrdlicka, Harvard should have recognized that its assets did not match its liabilities. In technical terms the duration of its liabilities was shorter than the duration of its assets.

Harvard faced five choices:

1. Liquidate a portion of the endowment

   But a lot of the endowment is illiquid and cannot be sold.

2. Cut expenses

   Universities are like government bureaucracies: big, bloated and inefficient. You can hardly fire anyone. So there is a limit to how much can be cut.

3. Increase donations

   It's embarrassing to ask for funds to replace those lost as a result of mismanagement.

4. Increase other revenue

   Harvard could rescind its need-blind financial-aid policy. But it turns out this doesn't save much money.

5. Borrow

Harvard did (5). It issued $2.5 billion in bonds and more than doubled its leverage ratio between 2008 and 2009. It did try to cut expenses, and deferred its Allston campus expansion. Was the endowment a rainy day fund Joseph could use to save his family and all of Egypt, as suggested by Dershowitz? No. Harvard actually *reduced* its payout ratio in 2009, preferring to keep as much of the endowment intact as it could. Maybe Hansmann is right in suggesting that prestige maximization is the driving motivation in endowment management. After all, (almost) everyone likes to be well-endowed.

**Table 1**

| Asset Class | Typical Time between Transactions | Annualized Turnover |
|---|---|---|
| | | |
| Public Equities | Within seconds | Well over 100% |
| | | |
| OTC (Pinksheet) Equities | Within a day, but many stocks over a week | Approx 35% |
| | | |
| Corporate Bonds | Within a day | 25-35% |
| | | |
| Municipal Bonds | Approx 6 months, with 5% of muni bonds trading more infrequently than once per decade | Less than 10% |
| | | |
| Private Equity | Funds last for 10 years; the median investment duration is 4 years; secondary trade before exist is relatively unusual | Less than 10% |
| | | |
| Residential Housing | 4-5 years, but ranges from months to decades | Approx 5% |
| | | |
| Institutional Real Estate | 8-11 years | Approx 7% |
| | | |
| Institutional Infrastructure | 50-60 years for initial commitment, some as long as 99 years | Negligible |
| | | |
| Art | 40-70 years | Less than 15% |

**Figure 2**

**Panel A**

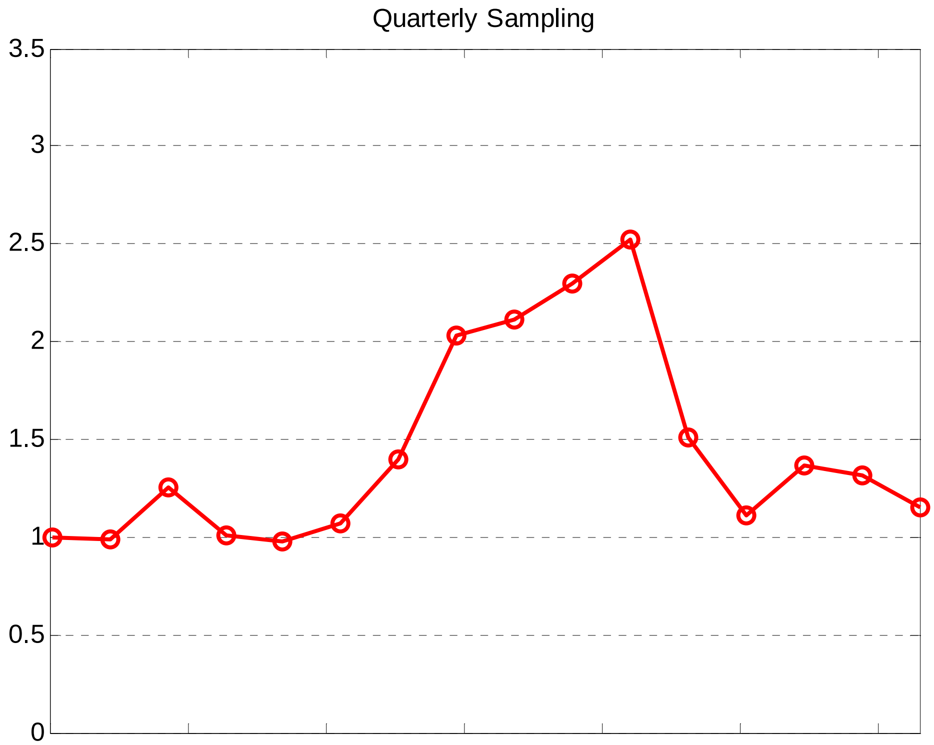


**Figure 2**

**Panel B**

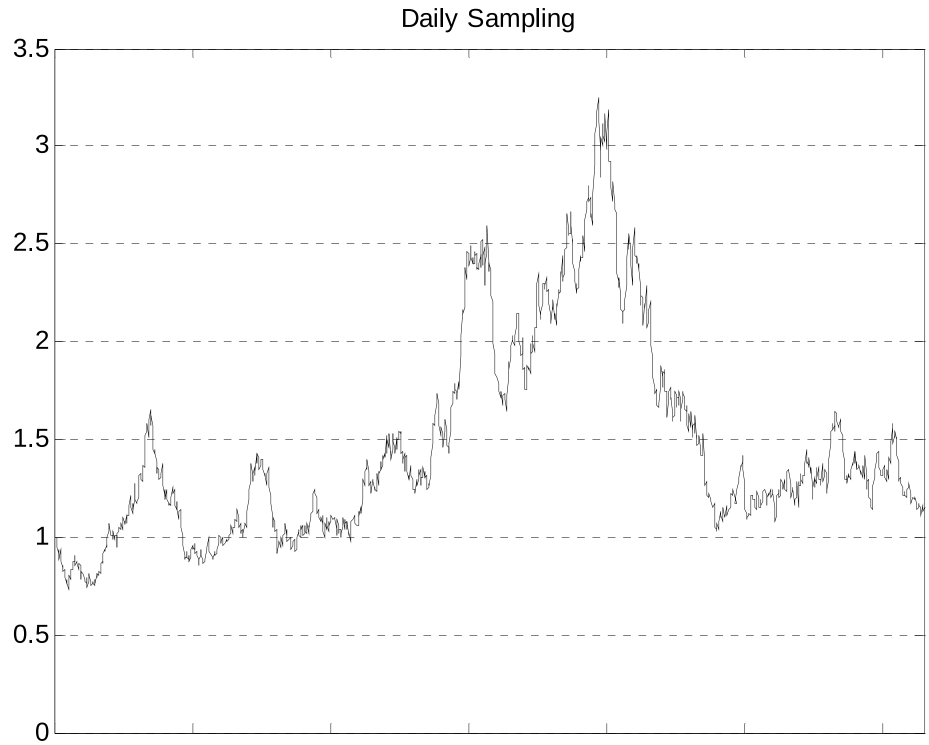


**Figure 2**

**Panel C**

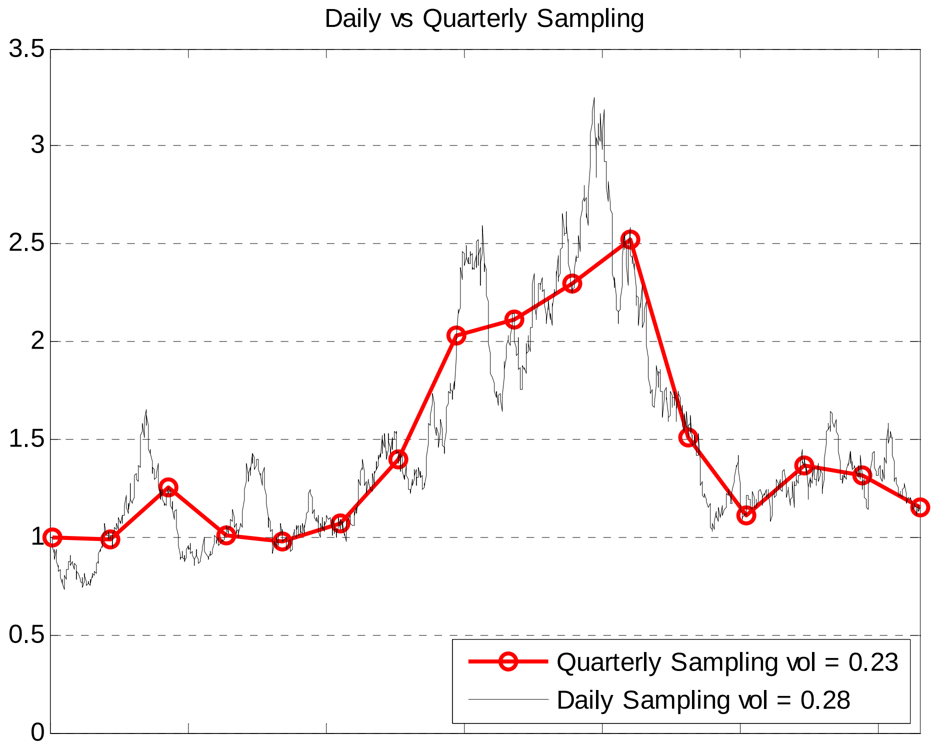


**Figure 3**

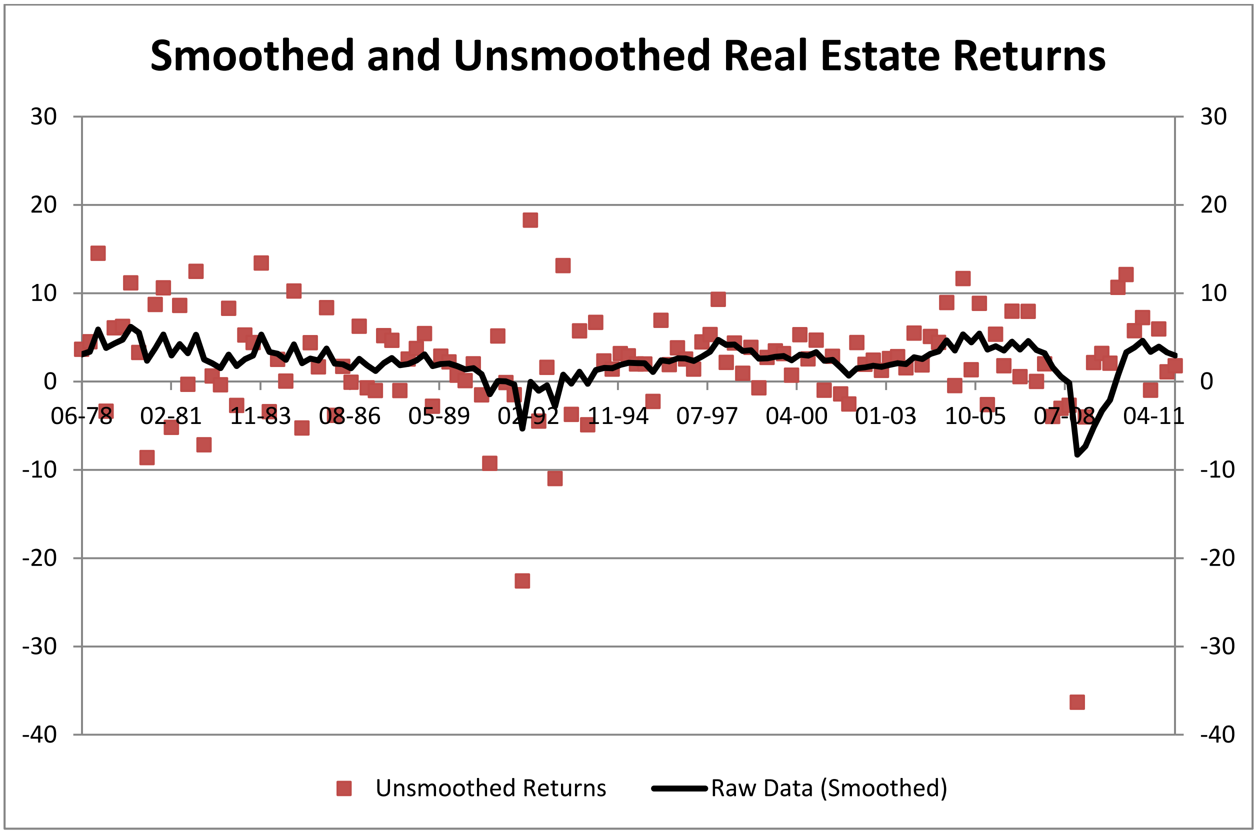

Smoothed and Unsmoothed Real Estate Returns
30
20
10
0
-10
-20
-30
-40
06-78
02-81
11-83
08-86
05-89
02-92
11-94
07-97
04-00
01-03
10-05
07-08
04-11
Unsmoothed Returns
Raw Data (Smoothed)

**Figure 4**

**Panel A**

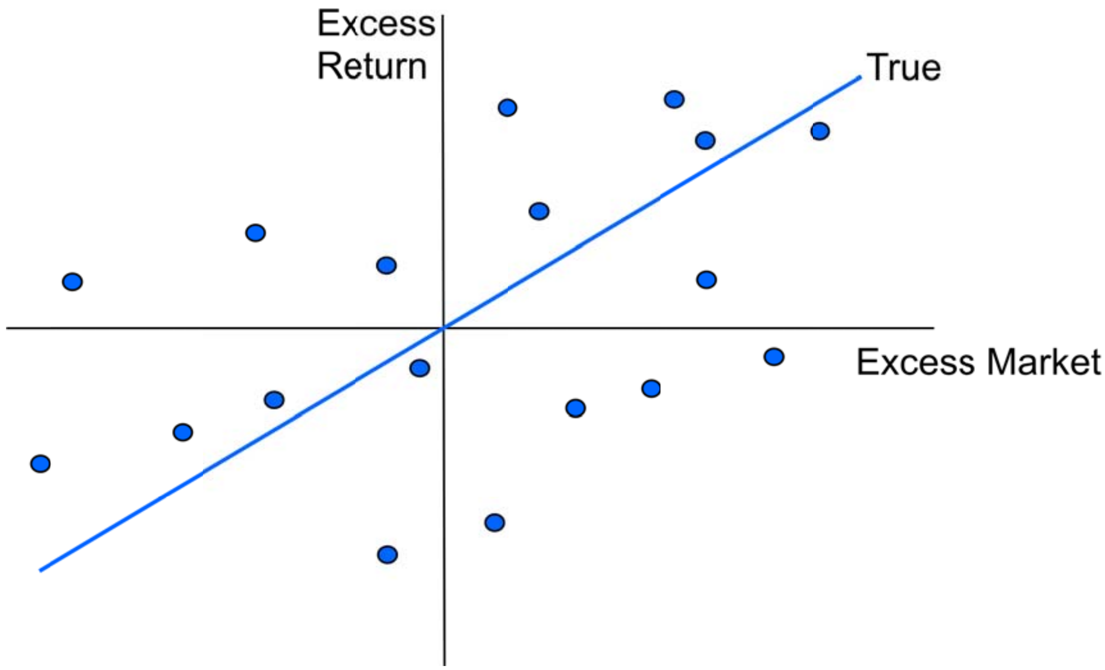


**Figure 4**

**Panel B**

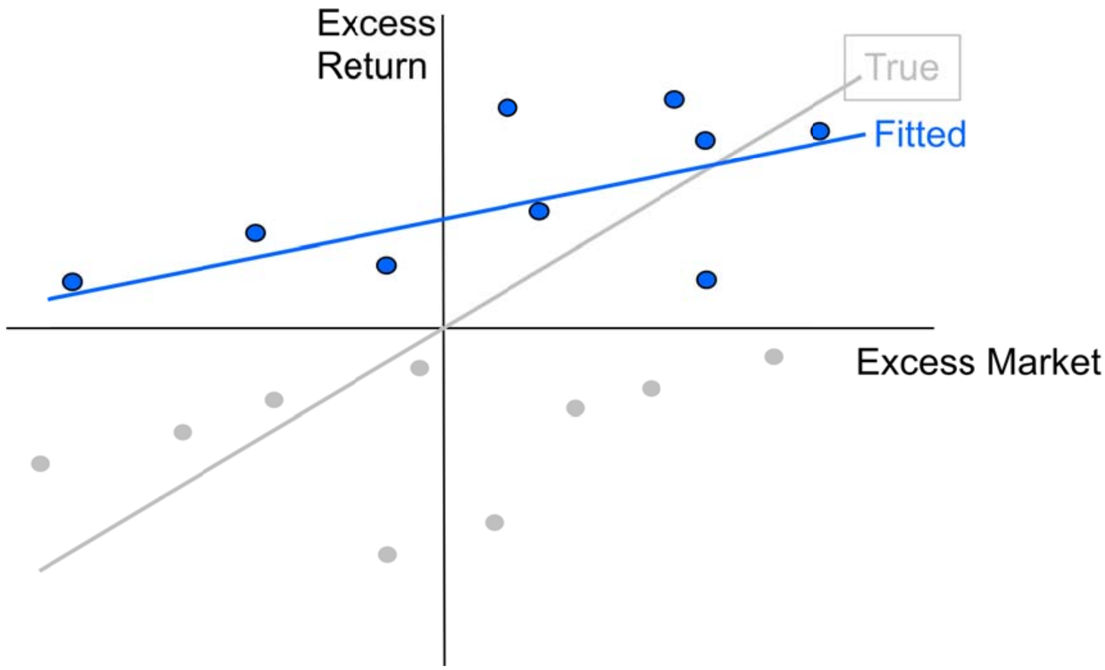
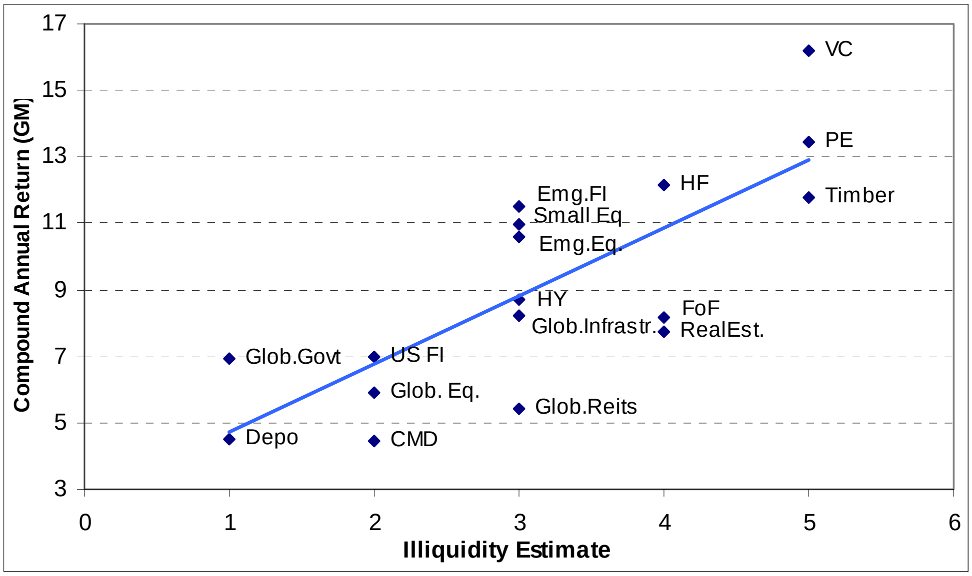


**Figure 5**




From Ilmanen (2011). Need to get permission from Wiley.

**Table 6**

**Panel A**

| Average Time Between Liquidity Events (or Average Turnover) | Optimal Rebalance Value |
|---|---|
| 10 years | 0.05 |
| 5 years | 0.11 |
| 2 years | 0.24 |
| 1 year | 0.37 |
| ½ year | 0.44 |
| Continuous Trading | 0.59 |

**Panel B**

| Average Time Between Liquidity Events (or Average Turnover) | Illiquidity Risk Premium |
|---|---|
| 10 years | 6.0% |
| 5 years | 4.3% |
| 2 years | 2.0% |
| 1 year | 0.9% |
| ½ year | 0.7% |
| Continuous Trading | 0.0% |