The Conditional Captital Asset Pricing Model Revisited: Evidence from High-Frequency Betas

MANAGEMENT SCIENCE 2020 Fabian Hollstein et. al.

苏锦华

2022.04

Content

- Introduction
- Data and Methodology
  - Historical Beta, Realized Beta, Hybric Beta
- The Conditional CAPM Revisited
  - Size, Value, Monmentum
  - Alternative Windows Lengths, Frequencies, Conditioning
- The Precision of High-Frequency Betas
  - RMSE, Regression
- Conclusion

Introduction

" 本文重新研究了条件CAPM的在高频数据下的表现。在一个2500多只股票的大样本，包括1996年至2014年期间的每日和高频数据，以实证检验条件资本资产定价模型是否能够解释资产价格异常。 "

Conditional CAPM能否解释超额收益

conditional CAPM(day): small, **big**, growth, value, **loser**, and **winner**

value (value minus growth) *momentum* (winners minus losers)
size (small minus big)

conditional CAPM(30m): **small**, **big**, **growth**, **value**, loser, and **winner**

Data and Methodology

Data

- 时间：1996年1月至2014年12月的样本期美国股票
- 来源:
  - 证券价格研究中心（CRSP）的每日和每月价格和回报以及股息支付和流通股数据
  - Tick History（TRTH）数据库收集高频价格数据
- 范围:
  - 在纽约证券交易所（NYSE）、美国证券交易所（AMEX）和美国证券交易商协会自动报价系统（NASDAQ）上交易的所有股票

- 其他辅助数据
  - 利率期限结构的数据--IvyDB
  - 无风险（一个月期国库券）利率--Kenneth French
- 数据处理
  - 排除了非常缺乏流动性的股票
  - 交易时间过滤
  - 利用CRSP提供的股票分割和分配数据来补充TRTH数据，以调整TRTH隔夜收益率
  - 为了防止小型股票非交易的潜在影响（Gorodnichenko和Weber 2016）主要研究30分钟的频率

Data and Methodology

Historical Beta

$$\beta_{j,t}^{\text{HIST}} = \frac{\text{cov}\left(r_j - r_f, r_M - r_f\right)}{\text{Var}\left(r_M - r_f\right)}$$

$$r_{j,\tau} - r_{f,\tau} = \alpha_{j,t} + \beta_{j,t}^{(0)}\left(r_{M,\tau} - r_{f,\tau}\right) + \beta_{j,t}^{(1)}\left(r_{M,\tau-1} - r_{f,\tau-1}\right)$$

$$+ \beta_{j,t}^{(2)}\left(\sum_{n=2}^{l} r_{M,\tau-n} - r_{f,\tau-n}\right) + \epsilon_{j,\tau}$$

Data and Methodology

Realized Beta

$$\beta_{j,t}^{\text{HF}} = \frac{\sum_{\tau=1}^{N} r_{j,\tau} r_{M,\tau}}{\sum_{\tau=1}^{N} r_{M,\tau}^2}$$

Hybird Beta

$$\beta_{j,t}^{\text{BV}} = \frac{\sigma_{j,t,T}^{\mathbb{Q}} \sum_{i=1}^{N} \omega_{i,t} \sigma_{i,t,T}^{\mathbb{Q}} \rho_{ji,t}^{\mathbb{Q}}}{\left(\sigma_{M,t,T}^{\mathbb{Q}}\right)^2}$$

Data and Methodology

Summary Statistics and Correlation Analysis

Table 1. Summary Statistics and Sample Correlations

	Summary statistics								Correlations									
	Number of observations	$Mean_{vw}$	Standard deviation	AR(1)	$q^{0.05}$	$q^{0.25}$	$q^{0.75}$	$q^{0.95}$	$HIST^{(1)}$	$HIST^{(2)}$	HF_{30}	$HF_{30}^{(1)}$	$HF_{30}^{(5)}$	HF_5	HF_{75}	HF_{130}	BV	BV_{30}
HIST	221,842	1.00	0.37	0.97	0.44	0.73	1.27	1.64	0.88	0.80	0.78	0.78	0.73	0.74	0.79	0.80	0.74	0.67
$HIST^{(1)}$	221,842	1.02	0.42	0.96	0.44	0.77	1.34	1.80	—	0.92	0.70	0.73	0.74	0.65	0.72	0.74	0.69	0.64
$HIST^{(2)}$	221,842	1.02	0.44	0.95	0.41	0.77	1.37	1.87	—	—	0.65	0.68	0.69	0.60	0.67	0.68	0.65	0.61
HF_{30}	221,842	0.95	0.36	0.97	0.32	0.57	1.11	1.46	—	—	—	0.92	0.79	0.95	0.95	0.92	0.72	0.78
$HF_{30}^{(1)}$	221,842	0.99	0.37	0.95	0.40	0.68	1.20	1.60	—	—	—	—	0.90	0.83	0.92	0.91	0.73	0.73
$HF_{30}^{(5)}$	221,842	1.02	0.40	0.93	0.43	0.75	1.28	1.72	—	—	—	—	—	0.70	0.82	0.84	0.71	0.66
HF_5	221,842	0.93	0.35	0.97	0.26	0.47	1.01	1.33	—	—	—	—	—	—	0.89	0.86	0.67	0.74
HF_{75}	221,842	0.96	0.37	0.96	0.35	0.62	1.15	1.52	—	—	—	—	—	—	—	0.95	0.74	0.77
HF_{130}	221,842	0.96	0.37	0.95	0.36	0.64	1.16	1.54	—	—	—	—	—	—	—	—	0.76	0.75
BV	75,548	0.99	0.24	0.88	0.68	0.88	1.19	1.46	—	—	—	—	—	—	—	—	—	0.89
BV_{30}	73,883	1.00	0.22	0.90	0.72	0.88	1.17	1.42	—	—	—	—	—	—	—	—	—	—

Notes. This table presents summary statistics of various beta estimates used in the paper. For all HIST estimators and BV, we use a sampling window of 12 months, whereas for all HF estimators and BV_{30}, we use a 6-month window. The summary statistics are similar for alternative sampling windows. Additionally, we present averages of firm-level correlations of the different beta estimation techniques. The total number of monthly estimates for individual firms is indicated by the number of observations. $Mean_{vw}$ is the overall value-weighted average of the estimates over the entire sample period. Standard deviation and AR(1) present further summary statistics on the average of firm-level standard deviations and first-order autocorrelations, whereas $q^{0.05}$, $q^{0.25}$, $q^{0.75}$, and $q^{0.95}$ provide the respective average quantiles. The sample spans January 1996 until December 2014.

- 所有历史和高频估计器的总观测数都要高得多，主要是因为混合估计器只能在每个时间点对标准普尔500指数
- 添加滞后Beta会显著增加HIST中的测量误差,估计器中的噪声自然会增加标准偏差，并在AR（1）系数中产生衰减偏差
- 由于HIST的加权平均值为1.00，几乎没有证据表明HIST存在系统性偏差,因此，我们主要讨论无滞后调整的估计量
- 相比之下，高频估计器的加权平均β值略低于1，对于某些股票而言，不频繁且不同步的交易可能会使贝塔系数偏向于零。加上一个滞后，加权平均值上升
- 观察到历史每日和30分钟高频估计器之间的相关性为0.78

The Conditional CAPM Revisited

Setup

实证试图研究以下问题：高频贝塔估计值下条件CAPM表现如何？高频贝塔能帮助解释为什么小公司的股市收益通常优于大公司吗？如何解释高频Beta相对于低频Beta的优势？条件CAPM模型能否解释为什么过去的赢家股票通常表现优于过去的输家股票？

本文遵循Lewellen和Nagel（2006）提出的方法，直接估计条件β，同时获得条件α的估计。借此对整个样本进行划分(季度非重叠窗口)，容许条件β时序变化，用于无条件测试。对于每个异常投资组合的条件α和β。得到条件α的时间序列，我们可以通过测试平均α是否显著不同于零来直接测试条件CAPM。使用稳健的Newey和West（1987）标准误差和四个滞后来评估统计显著性。

对于所有异常投资组合，利用纽约证券交易所断点，对相应股票条件α和β进行加权平均得到投资组合的条件α和β。我们按照Lewellen和Nagel（2006）的观点构建了Size和Value投资组合。构造的出发点主要是账面市盈率相关数据，每个月将这些股票独立分类为25个大小的账面市值投资组合，并根据这些投资组合的组合进行测试：

- S(small)是五个低市值投资组合的平均值
- B(big)是五个高市值投资组合的平均值，而S–B是它们的差异
- G(groth)是五个低账面市值投资组合的平均值
- V(value)是五个高账面市值投资组合的平均值，V–G是价值和增长投资组合之间的差异

- momentum是根据之前12个月的回报率将股票分为10个投资组合，而跳过最近1个月（Jegadeesh和Titman 1993）
- L（loser）和W（winner）是底部和顶部的10%位，W-L是赢家和输家投资组合之间的差异

The Conditional CAPM Revisited

Unconditional & Conditional CAPM

Table 2. Conditional CAPM Tests

	Size				Value				Momentum			
	α_S	α_B	α_{S-B}	β_{S-B}	α_G	α_V	α_{V-G}	β_{V-G}	α_L	α_W	α_{W-L}	β_{W-L}
Unconditional	0.0065**	0.0001	0.0063**	0.2303***	0.0046**	0.0042	−0.0004	−0.1679*	−0.0057	0.0049*	0.0106**	−0.5550***
	(0.0032)	(0.0011)	(0.0028)	(0.0785)	(0.0019)	(0.0027)	(0.0027)	(0.0970)	(0.0037)	(0.0028)	(0.0053)	(0.2102)
Daily data												
HIST	0.0079**	0.0010	0.0069**	0.0620	0.0060**	0.0052**	−0.0007	−0.1142*	0.0009	0.0025	0.0016	0.0775
	(0.0035)	(0.0009)	(0.0030)	(0.0860)	(0.0023)	(0.0024)	(0.0021)	(0.0677)	(0.0029)	(0.0025)	(0.0043)	(0.1599)
HIST[1]	0.0074**	0.0012	0.0061*	0.1144	0.0050**	0.0053**	0.0002	−0.1404*	0.0003	0.0023	0.0021	0.0357
	(0.0036)	(0.0009)	(0.0031)	(0.0806)	(0.0023)	(0.0025)	(0.0021)	(0.0756)	(0.0029)	(0.0025)	(0.0044)	(0.1839)
HIST[2]	0.0078*	0.0013	0.0065*	0.1376	0.0051**	0.0057**	0.0006	−0.1231*	−0.0002	0.0025	0.0027	−0.0310
	(0.0041)	(0.0010)	(0.0035)	(0.0836)	(0.0024)	(0.0028)	(0.0021)	(0.0729)	(0.0034)	(0.0024)	(0.0047)	(0.1918)
High-frequency data												
HF_{30}	0.0013	−0.0006	0.0019	−0.1609**	0.0001	0.0013	0.0012	−0.1110**	−0.0086*	−0.0038	0.0048	0.0942
	(0.0031)	(0.0010)	(0.0029)	(0.0805)	(0.0015)	(0.0024)	(0.0022)	(0.0554)	(0.0046)	(0.0027)	(0.0050)	(0.1311)
$HF_{30}^{(1)}$	0.0004	−0.0010	0.0014	−0.0392	−0.0007	0.0008	0.0015	−0.1306**	−0.0094**	−0.0040	0.0054	0.1037
	(0.0031)	(0.0011)	(0.0029)	(0.0865)	(0.0015)	(0.0024)	(0.0022)	(0.0618)	(0.0046)	(0.0025)	(0.0049)	(0.1370)
$HF_{30}^{(5)}$	0.0000	−0.0010	0.0010	0.0267	−0.0008	0.0009	0.0017	−0.1414**	−0.0090**	−0.0041*	0.0050	0.0980
	(0.0034)	(0.0011)	(0.0030)	(0.0809)	(0.0017)	(0.0024)	(0.0021)	(0.0667)	(0.0042)	(0.0024)	(0.0047)	(0.1469)

Notes. This table reports average conditional alphas for different value-weighted anomaly portfolios. We present the results for size, value, and momentum portfolios. For size and value, we build 25 independently sorted portfolios. S is the average of the five low-market-cap portfolios, and B is the average of the five high-market-cap portfolios, whereas S–B is their difference. Similarly, G is the average of the five low-book-to-market portfolios, and V is the average of the five high-book-to-market portfolios. V–G is the difference between the value and growth portfolios. For momentum, we sort the stocks into 10 portfolios according to their return over the past 12 months (skipping 1 month). L and W are the bottom and top deciles, whereas W–L is their difference. Unconditional denotes the unconditional CAPM based on monthly returns for the full

The Conditional CAPM Revisited

Underconditioning Bias

$$\alpha_j^{\text{unc}} - \bar{\alpha}_j^{\text{cond}} = \left(1 + \frac{\bar{R}_M^2}{\sigma_M^2} \right) \text{cov} \left(\beta_{j,t}, \mathbb{E}_t \left(R_{M,t} \right) \right)$$

$$\underbrace{- \left(\frac{\bar{R}_M}{\sigma_M^2} \right) \text{cov} \left(\beta_{j,t}, \mathbb{E}_t \left(R_{M,t}^2 \right) \right)}_{\text{market-timing bias}}$$

$$\underbrace{- \left(\frac{\bar{\bar{R}}_M}{\sigma_M^2} \right) \text{cov} \left(\beta_{j,t}, \sigma_{M,t}^2 \right),}_{\text{volatility-timing bias}}$$

继Lewellen和Nagel（2006年）和Boguth等人（2011年）之后，我们现在研究了欠条件(underconditioning)可能引入的潜在偏差（即当β确实随时间变化时，使用无条件CAPM而不是传统CAPM）。Boguth等人（2011年）表明，如果条件α与市场超额收益不相关，则无条件α中的偏差等于无条件α与平均条件α之间的差异。

欠条件偏差包括市场时间偏差和波动时间偏差。市场时机偏差源于条件β与预期市场超额收益或预期市场超额收益平方的协变量，反映了β可能随商业周期而变化的事实（Jagannathan和Wang，1996），"通常小于0.20%"的偏差，"最大为0.35%"。波动时间偏差源于条件β与市场条件波动的负协变量，这似乎对动量尤其明显，波动时间偏差达到每月0.10%到1.01%之间的值。对于本文样本中的投资组合，Alpha的绝对值在大盘对开本的每月0.01%和momentum反常投资组合的每月1.06%之间。

- 因此，我们观察到的无条件回报的幅度可能与欠条件的偏差一致。

The Conditional CAPM Revisited

Alternative Windows Lengths, Frequencies, Conditioning

Table 3. Conditional CAPM Tests: Different Window Lengths

	Size				Value				Momentum			
	α_S	α_B	α_{S-B}	β_{S-B}	α_G	α_V	α_{V-G}	β_{V-G}	α_L	α_W	α_{W-L}	β_{W-L}
Unconditional	0.0065**	0.0001	0.0063**	0.2303***	0.0046**	0.0042	−0.0004	−0.1679*	−0.0057	0.0049*	0.0106**	−0.5550***
	(0.0032)	(0.0011)	(0.0028)	(0.0785)	(0.0019)	(0.0027)	(0.0027)	(0.0970)	(0.0037)	(0.0028)	(0.0053)	(0.2102)
Monthly												
HIST	0.0081***	0.0009	0.0073**	0.0678	0.0065***	0.0057**	−0.0009	−0.1102**	0.0009	0.0071***	0.0062	0.0300
	(0.0031)	(0.0008)	(0.0028)	(0.0550)	(0.0019)	(0.0023)	(0.0023)	(0.0427)	(0.0039)	(0.0024)	(0.0054)	(0.1137)
HF_{30}	0.0017	−0.0002	0.0019	−0.1655***	0.0006	0.0021	0.0015	−0.1105***	−0.0083**	0.0001	0.0084	0.0462
	(0.0031)	(0.0009)	(0.0030)	(0.0493)	(0.0018)	(0.0023)	(0.0023)	(0.0355)	(0.0038)	(0.0029)	(0.0052)	(0.0965)
$HF_{30}^{(1)}$	0.0010	−0.0008	0.0018	−0.0441	−0.0001	0.0016	0.0017	−0.1317***	−0.0093**	0.0001	0.0093*	0.0607
	(0.0030)	(0.0010)	(0.0029)	(0.0535)	(0.0016)	(0.0022)	(0.0022)	(0.0397)	(0.0037)	(0.0029)	(0.0051)	(0.1009)
Semiannually												
HIST	0.0083**	0.0010	0.0074**	0.0725	0.0059**	0.0052**	−0.0006	−0.1123	0.0047	0.0024	−0.0023	0.1037
	(0.0038)	(0.0009)	(0.0033)	(0.1162)	(0.0028)	(0.0023)	(0.0022)	(0.0862)	(0.0032)	(0.0022)	(0.0033)	(0.1472)
HF_{30}	0.0016	−0.0004	0.0021	−0.1517	0.0002	0.0016	0.0014	−0.1069	−0.0043	−0.0029	0.0014	0.1240
	(0.0028)	(0.0012)	(0.0027)	(0.1109)	(0.0014)	(0.0021)	(0.0022)	(0.0721)	(0.0043)	(0.0023)	(0.0036)	(0.1248)
$HF_{30}^{(1)}$	0.0007	−0.0009	0.0016	−0.0320	−0.0005	0.0011	0.0016	−0.1250	−0.0046	−0.0031	0.0015	0.1304
	(0.0030)	(0.0013)	(0.0029)	(0.1182)	(0.0016)	(0.0022)	(0.0022)	(0.0805)	(0.0043)	(0.0020)	(0.0035)	(0.1297)
Annually												
HIST	0.0083**	0.0010	0.0073**	0.0735	0.0072**	0.0044*	−0.0028	−0.0871	0.0057	0.0027	−0.0030	0.0581
	(0.0036)	(0.0007)	(0.0031)	(0.1430)	(0.0030)	(0.0022)	(0.0022)	(0.1119)	(0.0036)	(0.0018)	(0.0030)	(0.1458)
HF_{30}	0.0020	−0.0006	0.0026	−0.1452	0.0018	0.0010	−0.0009	−0.0906	−0.0020	−0.0033	−0.0013	0.0866
	(0.0028)	(0.0010)	(0.0029)	(0.1434)	(0.0015)	(0.0022)	(0.0019)	(0.0936)	(0.0027)	(0.0023)	(0.0025)	(0.1163)
$HF_{30}^{(1)}$	0.0011	−0.0010	0.0020	−0.0243	0.0012	0.0005	−0.0008	−0.1004	−0.0026	−0.0036	−0.0010	0.0876

使用三个月的窗口来估计条件α和β。然而，窗口长度涉及到一个权衡：一方面，较短的窗口可能会提高模型的性能，因为它允许在beta中有更大的变化；另一方面，更长的窗口可以提高条件alpha和beta估计的精度，并以这种方式改善模型性能。因此，在表3中检查了1个月、6个月和12个月的替代窗口长度。只报告主要估计器HIST、HF30和HF（1）30的结果。如表所示，**这些变化并不影响主要结论。**

The Conditional CAPM Revisited

Alternative Windows Lengths, Frequencies, Conditioning

Table 4. Conditional CAPM Tests: Different Frequencies

	Size				Value				Momentum			
	α_S	α_B	α_{S-B}	β_{S-B}	α_G	α_V	α_{V-G}	β_{V-G}	α_L	α_W	α_{W-L}	β_{W-L}
Unconditional	0.0065**	0.0001	0.0063**	0.2303***	0.0046**	0.0042	−0.0004	−0.1679*	−0.0057	0.0049*	0.0106**	−0.5550***
	(0.0032)	(0.0011)	(0.0028)	(0.0785)	(0.0019)	(0.0027)	(0.0027)	(0.0970)	(0.0037)	(0.0028)	(0.0053)	(0.2102)
Different HF estimators												
HF$_5$	0.0023	−0.0005	0.0028	−0.3182***	0.0007	0.0016	0.0009	−0.0832	−0.0087*	−0.0036	0.0052	0.0791
	(0.0033)	(0.0011)	(0.0029)	(0.0707)	(0.0016)	(0.0025)	(0.0023)	(0.0505)	(0.0049)	(0.0027)	(0.0051)	(0.1279)
HF$_5^{(10)}$	0.0001	−0.0012	0.0013	−0.0201	−0.0008	0.0007	0.0014	−0.1370**	−0.0096**	−0.0038	0.0057	0.0869
	(0.0032)	(0.0011)	(0.0029)	(0.0861)	(0.0016)	(0.0024)	(0.0022)	(0.0632)	(0.0045)	(0.0025)	(0.0048)	(0.1400)
HF$_{75}$	0.0006	−0.0005	0.0011	−0.0840	−0.0002	0.0012	0.0013	−0.1269**	−0.0090*	−0.0037	0.0053	0.1026
	(0.0031)	(0.0010)	(0.0029)	(0.0809)	(0.0015)	(0.0023)	(0.0022)	(0.0593)	(0.0047)	(0.0026)	(0.0050)	(0.1344)
HF$_{75}^{(3)}$	−0.0005	−0.0008	0.0003	0.0526	−0.0009	0.0007	0.0016	−0.1421*	−0.0092**	−0.0040	0.0052	0.1011
	(0.0033)	(0.0010)	(0.0031)	(0.0807)	(0.0018)	(0.0023)	(0.0022)	(0.0714)	(0.0039)	(0.0025)	(0.0046)	(0.1609)
HF$_{130}$	0.0010	−0.0007	0.0017	−0.0541	−0.0001	0.0012	0.0013	−0.1244**	−0.0085*	−0.0038	0.0047	0.0913
	(0.0032)	(0.0011)	(0.0029)	(0.0781)	(0.0016)	(0.0024)	(0.0022)	(0.0614)	(0.0046)	(0.0027)	(0.0050)	(0.1354)
HF$_{130}^{(3)}$	0.0002	−0.0003	0.0005	0.0876	−0.0006	0.0014	0.0020	−0.1482*	−0.0086**	−0.0037	0.0049	0.0806
	(0.0035)	(0.0012)	(0.0031)	(0.0822)	(0.0020)	(0.0023)	(0.0022)	(0.0796)	(0.0037)	(0.0024)	(0.0044)	(0.1801)

Notes. This table reports average conditional alphas for different value-weighted anomaly portfolios. We present the results for size, value, and momentum portfolios. For size and value, we build 25 independently sorted portfolios. S is the average of the five low-market-cap portfolios, and B is the average of the five high-market-cap portfolios, whereas S–B is their difference. Similarly, G is the average of the five low-book-to-market portfolios, and V is the average of the five high-book-to-market portfolios. V–G is the difference between the value and growth portfolios. For momentum, we sort the stocks into 10 portfolios according to their return over the past 12 months (skipping 1 month). L and W are the bottom

检查高频估计模型的替代采样频率对结果的影响。根据有无滞后的5分钟、75分钟和130分钟数据检验了参数估计。对于所有这些估计模型，我们获得了与HF30定性相似的结果：基于高频数据的条件CAPM可以解释Size、Value和Momentum异常。

The Conditional CAPM Revisited

Alternative Windows Lengths, Frequencies, Conditioning

Table 5. Conditional CAPM Tests: Conditioning Approaches

	Size				Value				Momentum			
	α_S	α_B	α_{S-B}	β_{S-B}	α_G	α_V	α_{V-G}	β_{V-G}	α_L	α_W	α_{W-L}	β_{W-L}
Unconditional	0.0065**	0.0001	0.0063**	0.2303***	0.0046**	0.0042	−0.0004	−0.1679*	−0.0057	0.0049*	0.0106**	−0.5550***
	(0.0032)	(0.0011)	(0.0028)	(0.0785)	(0.0019)	(0.0027)	(0.0027)	(0.0970)	(0.0037)	(0.0028)	(0.0053)	(0.2102)
Lagged proxies												
HIST	0.0077**	0.0004	0.0073**	0.0515	0.0059**	0.0044*	−0.0015	−0.1244**	0.0008	0.0011	0.0002	0.0058
	(0.0034)	(0.0009)	(0.0031)	(0.0847)	(0.0023)	(0.0024)	(0.0023)	(0.0717)	(0.0030)	(0.0026)	(0.0046)	(0.1499)
HF_{30}	0.0012	−0.0014	0.0025	−0.1704**	0.0000	0.0007	0.0007	−0.1139*	−0.0093**	−0.0049	0.0044	0.0357
	(0.0032)	(0.0012)	(0.0029)	(0.0799)	(0.0016)	(0.0024)	(0.0023)	(0.0593)	(0.0045)	(0.0030)	(0.0048)	(0.1267)
$HF_{30}^{(1)}$	−0.0000	−0.0019	0.0018	−0.0466	−0.0006	−0.0001	0.0006	−0.1318*	−0.0096**	−0.0052*	0.0044	0.0494
	(0.0032)	(0.0013)	(0.0030)	(0.0868)	(0.0016)	(0.0024)	(0.0023)	(0.0663)	(0.0046)	(0.0030)	(0.0048)	(0.1325)
Instruments: Full sample												
HIST	0.0078**	0.0005	0.0073**	0.0580	0.0060***	0.0045*	−0.0016	−0.1157*	0.0007	0.0009	0.0002	0.0714
	(0.0035)	(0.0009)	(0.0031)	(0.0772)	(0.0023)	(0.0024)	(0.0023)	(0.0687)	(0.0030)	(0.0026)	(0.0046)	(0.1287)
HF_{30}	0.0011	−0.0012	0.0023	−0.1608**	0.0001	0.0008	0.0007	−0.1146*	−0.0096**	−0.0051*	0.0045	0.0964
	(0.0032)	(0.0012)	(0.0029)	(0.0788)	(0.0016)	(0.0024)	(0.0023)	(0.0580)	(0.0046)	(0.0030)	(0.0051)	(0.1067)
$HF_{30}^{(1)}$	0.0008	−0.0013	0.0021	−0.1600*	0.0000	0.0004	0.0004	−0.1140*	−0.0097**	−0.0051*	0.0047	0.0916
	(0.0032)	(0.0013)	(0.0029)	(0.0839)	(0.0016)	(0.0025)	(0.0023)	(0.0643)	(0.0048)	(0.0030)	(0.0051)	(0.1070)
Instruments: Expanding window												
HIST	0.0091**	0.0006	0.0085**	0.1256*	0.0070***	0.0055**	−0.0015	−0.0450	0.0008	−0.0001	−0.0009	0.0333
	(0.0038)	(0.0010)	(0.0034)	(0.0743)	(0.0025)	(0.0027)	(0.0024)	(0.0708)	(0.0035)	(0.0029)	(0.0053)	(0.1371)
HF_{30}	0.0012	−0.0014	0.0026	−0.0633	0.0005	0.0008	0.0003	−0.0577	−0.0096*	−0.0066**	0.0031	0.0488
	(0.0033)	(0.0013)	(0.0032)	(0.0757)	(0.0018)	(0.0026)	(0.0024)	(0.0618)	(0.0054)	(0.0033)	(0.0058)	(0.1236)
$HF_{30}^{(1)}$	0.0011	−0.0013	0.0024	−0.0658	0.0006	0.0006	−0.0000	−0.0516	−0.0098*	−0.0067*	0.0032	0.0440
	(0.0034)	(0.0014)	(0.0032)	(0.0773)	(0.0018)	(0.0028)	(0.0025)	(0.0642)	(0.0057)	(0.0033)	(0.0057)	(0.1228)

Boguth等人（2011年）警告说，可能会在分析中引入偏差，因为用于计算条件α的贝塔估计值在投资者进行投资之前是无法事先获得的。为了检查我们的结果是否受到这种偏见的影响，我们使用三种不同的替代方法重复前面的分析。首先，我们只是使用滞后的beta作为未来beta的代理。其次，我们估计下述回归方程。

$$\beta_{j,t} = a_j + b_j \beta_{j,t-1} + \epsilon_{j,t}$$

本文使用了这种方法的两种变体：一种是使用Boguth等人（2011）中的完整样本，另一种实时实现的，只使用时间t之前的数据，并使用36个月窗口初始化的扩展窗口。我们在表5中给出了这些方法的结果。结果与baseline非常相似。我们发现，基于高频数据的条件CAPM仍然可以解释所有三种情况下的大小、值和动量异常。

The Precision of High-Frequency Betas

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left(\beta_{t,T}^{\text{R}} - \beta_t \right)^2},$$

计算不同方法的样本外预测误差。每个月，我们都会根据过去12个月的估值，将这些股票分为五个投资组合。我们获得每个投资组合和方法的价值加权平均贝塔值，并使用随后一个月的30分钟回报率衡量每个投资组合的已实现贝塔值。第一行报告所有投资组合的平均RMSE。括号中数字的绝对值表示差异显著的投资组合份额为5%（例如，0.4表示差异显著的投资组合份额为40%）。用改进的Diebold–Mariano和Wilcoxon符号秩检验对上下三角矩阵进行显著性检验。括号中数字的符号表示显著差异的方向。

The Precision of High-Frequency Betas

	HIST	HIST$^{(1)}$	HF$_{30}$	HF$_{30}^{(1)}$
Panel A: Realized beta				
Average RMSE	0.1352	0.1542	*0.0949*	0.1180
HIST	—	−0.0190	0.0403	0.0172
	—	(−1.00)	(1.00)	(0.40)
HIST$^{(1)}$	0.0118	—	0.0593	0.0362
	(1.00)	—	(1.00)	(1.00)
HF$_{30}$	−0.0277	−0.0395	—	−0.0231
	(−1.00)	(−1.00)	—	(−1.00)
HF$_{30}^{(1)}$	−0.0064	−0.0181	0.0214	—
	(−0.40)	(−0.80)	(1.00)	—
Panel B: Realized beta with lag-adjustment				
Average RMSE	0.1491	0.1587	0.1344	*0.1276*
HIST	—	−0.0095	0.0147	0.0216
	—	(−0.40)	(0.20)	(0.80)
HIST$^{(1)}$	0.0060	—	0.0242	0.0311
	(0.40)	—	(0.40)	(1.00)
HF$_{30}$	−0.0011	−0.0070	—	0.0069
	(−0.20)	(−0.20)	—	(0.60)
HF$_{30}^{(1)}$	−0.0111	−0.0170	−0.0100	—
	(−0.60)	(−0.80)	(−0.60)	—
Panel C: Daily realized beta				
Average RMSE	0.1510	0.1561	0.1493	*0.1466*
HIST	—	−0.0052	0.0017	0.0044
	—	(−0.20)	(0.00)	(0.00)
HIST$^{(1)}$	0.0061	—	0.0069	0.0096
	(0.20)	—	(0.20)	(0.00)
HF$_{30}$	0.0058	−0.0003	—	0.0027
	(0.40)	(0.20)	—	(0.20)
HF$_{30}^{(1)}$	0.0023	−0.0039	−0.0035	—
	(0.00)	(−0.20)	(0.00)	—

Notes. This table presents the out-of-sample prediction errors of different approaches. Each month, we sort the stocks into five portfolios according to their estimate for HIST 12 months before. We obtain value-weighted average betas for each portfolio and approach and measure the realized beta of each portfolio using 30-minute returns over the subsequent month. The first row reports the average (Avg.) RMSE across all portfolios. The lowest average error is indicated by *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrices report the differences in RMSE, averaged over all portfolios. Similarly, the lower triangular matrices report the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of portfolios for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant for 40% of the portfolios). Figures printed in **bold** indicate that the differences are significant for *all* portfolios. We test significance with the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrices, respectively. The sign of the number in parentheses indicates the direction of the significant differences.
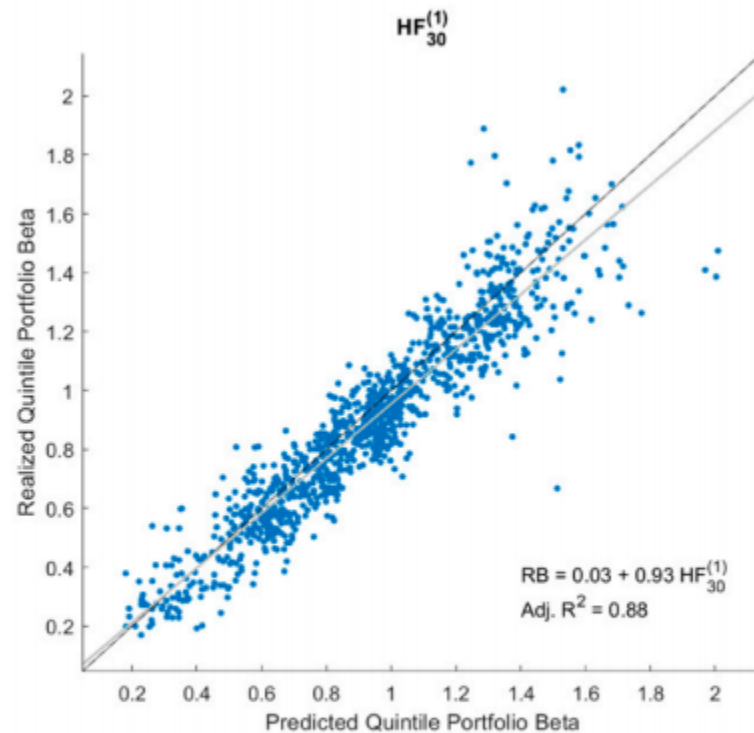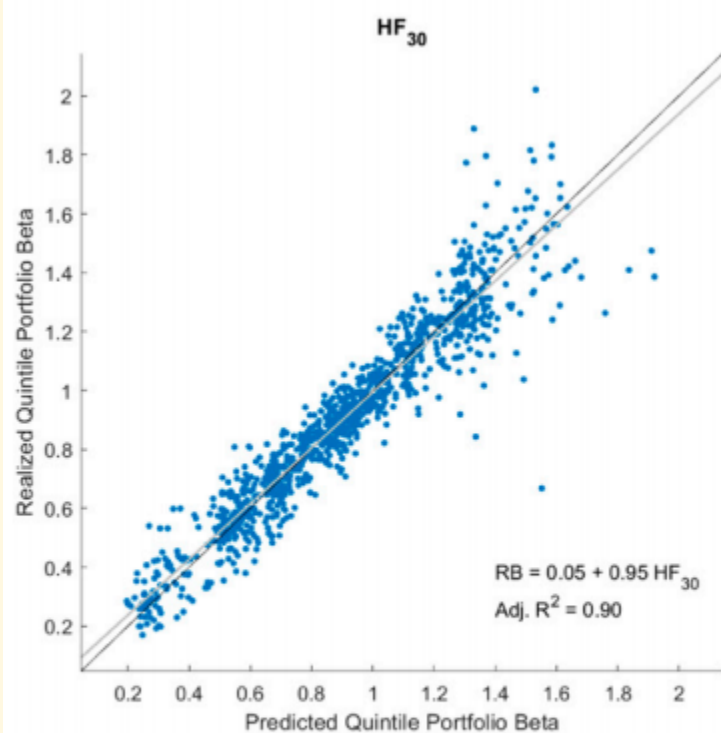
The Precision of High-Frequency Betas

Figure 1. (Color online) Predicted Beta vs. Realized Beta: Scatterplots

HIST

RB = 0.09 + 0.85 HIST
Adj. R^2 = 0.87

HIST$^{(1)}$

RB = 0.13 + 0.79 HIST$^{(1)}$
Adj. R^2 = 0.86
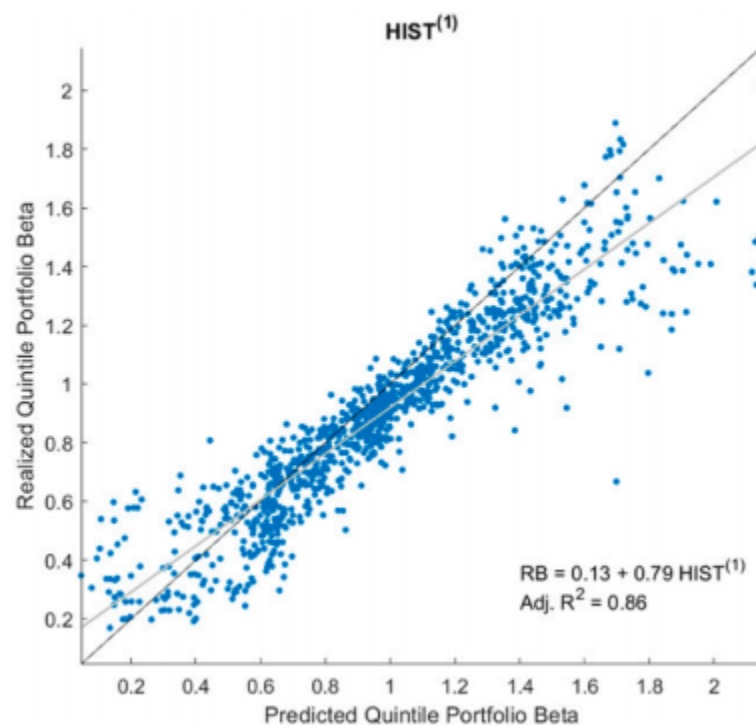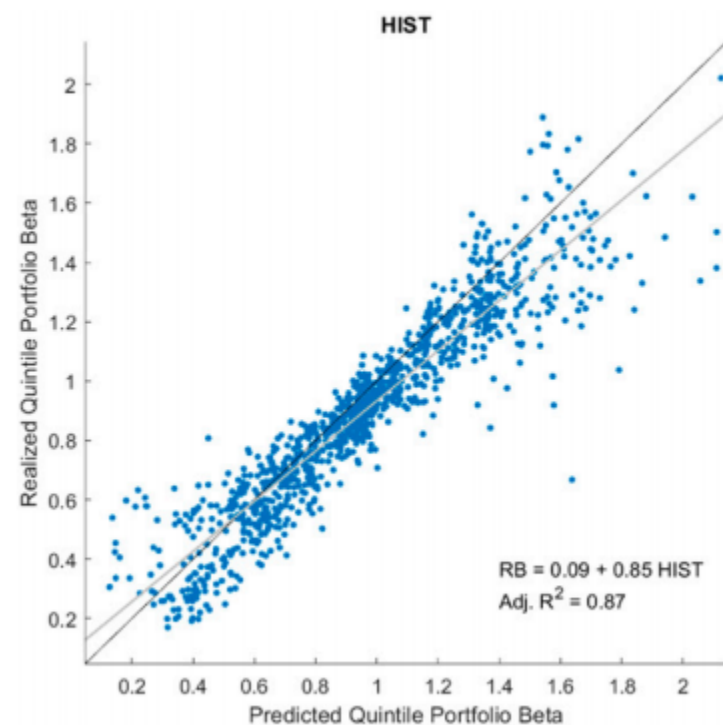
The Precision of High-Frequency Betas

Notes. This figure shows the scatterplots of predicted and realized beta. Each month, we sort the stocks into five portfolios according to their estimate for HIST 12 months before. For each of the portfolios, we compute the value-weighted average predicted beta as well as the realized beta over the next month, using 30-minute returns. We plot the realized quintile betas versus the predicted portfolio betas for each quintile, month, and estimation methodology. Each figure includes a 45° line and a fitted regression line, as well as the adjusted R^2 of the regression.

The Precision of High-Frequency Betas

Table 7. Prediction Errors: S&P 500 Stocks

	HIST	HIST$^{(1)}$	HF$_{30}$	HF$_{30}^{(1)}$	BV	BV$_{30}$
			Panel A: Realized beta			
Average RMSE	0.1320	0.1421	*0.0945*	0.1163	0.1072	0.1145
HIST	—	−0.0101	0.0375	0.0157	0.0247	0.0174
	—	(−0.60)	**(1.00)**	(0.40)	(0.60)	(0.40)
HIST$^{(1)}$	0.0107	—	0.0476	0.0258	0.0349	0.0275
	(0.80)	—	**(1.00)**	(0.80)	(0.60)	(0.40)
HF$_{30}$	−0.0215	−0.0322	—	−0.0218	−0.0128	−0.0201
	(−1.00)	**(−1.00)**	—	**(−1.00)**	(−0.40)	(−0.60)
HF$_{30}^{(1)}$	−0.0053	−0.0161	0.0162	—	0.0090	0.0017
	(−0.20)	(−0.80)	**(1.00)**	—	(0.20)	(0.00)
BV	−0.0114	−0.0221	0.0101	−0.0061	—	−0.0073
	(−0.40)	(−0.80)	(0.60)	(−0.20)	—	(−0.60)
BV$_{30}$	−0.0061	−0.0169	0.0154	−0.0008	0.0053	—
	(0.00)	(−0.20)	(0.60)	(−0.20)	(0.40)	—

The Precision of High-Frequency Betas

	Panel B: Realized beta with lag-adjustment					
Average RMSE	0.1511	0.1571	0.1301	*0.1292*	0.1294	0.1364
HIST	—	−0.0060	0.0211	0.0220	0.0218	0.0148
	—	(−0.40)	(0.60)	(0.60)	**(1.00)**	(0.60)
HIST$^{(1)}$	0.0040	—	0.0271	0.0280	0.0278	0.0208
	(0.40)	—	(0.80)	**(1.00)**	(0.80)	(0.60)
HF$_{30}$	−0.0089	−0.0129	—	0.0009	0.0007	−0.0063
	(−0.60)	(−0.40)	—	(0.00)	(0.00)	(0.00)
HF$_{30}^{(1)}$	−0.0116	−0.0156	−0.0027	—	−0.0002	−0.0072
	(−0.20)	(−0.80)	(0.00)	—	(0.00)	(−0.20)
BV	−0.0126	−0.0166	−0.0037	−0.0010	—	−0.0070
	(−0.60)	(−0.60)	(−0.20)	(0.20)	—	(−0.60)
BV$_{30}$	−0.0044	−0.0085	0.0044	0.0071	0.0082	—
	(−0.40)	(−0.60)	(0.00)	(0.20)	(0.60)	—

The Precision of High-Frequency Betas

			Panel C: Daily realized beta			
Average RMSE	0.1550	0.1579	0.1431	0.1476	*0.1416*	0.1486
HIST	—	−0.0029	0.0119	0.0074	0.0135	0.0065
	—	(−0.20)	(0.40)	(0.20)	(0.60)	(0.40)
HIST$^{(1)}$	0.0000	—	0.0148	0.0103	0.0164	0.0093
	(0.20)	—	(0.40)	(0.20)	(0.60)	(0.60)
HF$_{30}$	−0.0038	−0.0039	—	−0.0045	0.0015	−0.0055
	(−0.40)	(−0.40)	—	(0.00)	(0.20)	(0.00)
HF$_{30}^{(1)}$	−0.0027	−0.0027	0.0011	—	0.0061	−0.0010
	(0.20)	(0.00)	(0.20)	—	(0.20)	(0.00)
BV	−0.0116	−0.0116	−0.0077	−0.0089	—	−0.0070
	(−0.40)	(−0.40)	(−0.40)	(−0.60)	—	(−0.40)
BV$_{30}$	−0.0065	−0.0065	−0.0026	−0.0038	0.0051	—
	(0.00)	(−0.40)	(−0.20)	(−0.20)	(0.80)	—

Notes. This table presents the out-of-sample prediction errors of different approaches. Each month, we sort the stocks in the S&P 500 into five portfolios according to their estimate for HIST 12 months before. We obtain value-weighted average betas for each portfolio and approach and measure the realized beta of each portfolio using 30-minute returns over the subsequent month. The first row reports the average (Avg.) RMSE across all portfolios. The lowest average error is indicated by *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrices report the differences in RMSE, averaged over all portfolios. Similarly, the lower triangular matrices report the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of portfolios for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant for 40% of the portfolios). Figures printed in **bold** indicate that the differences *are* significant for *all* portfolios. We test significance with the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrices, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

The Precision of High-Frequency Betas

Table 8. Prediction Errors: Different Frequencies

	HIST	HIST$^{(1)}$	HF$_{30}$	HF$_{30}^{(1)}$	HF$_5$	HF$_5^{(10)}$	HF$_{75}$	HF$_{75}^{(1)}$	HF$_{130}$	HF$_{130}^{(1)}$
					Panel A: Realized beta					
Average RMSE	0.1353	0.1543	*0.0949*	0.1181	0.1012	0.1420	0.0980	0.1320	0.0995	0.1371
HIST	—	−0.0190	0.0404	0.0172	0.0341	−0.0067	0.0373	0.0033	0.0358	−0.0018
	—	**(−1.00)**	**(1.00)**	(0.40)	**(1.00)**	(−0.40)	**(1.00)**	(0.00)	**(1.00)**	(−0.20)
HIST$^{(1)}$	0.0120	—	0.0594	0.0362	0.0531	0.0122	0.0563	0.0223	0.0548	0.0172
	(1.00)	—	**(1.00)**	**(1.00)**	**(1.00)**	(0.20)	**(1.00)**	(0.60)	**(1.00)**	(0.40)
HF$_{30}$	−0.0273	−0.0393	—	−0.0232	−0.0063	−0.0472	−0.0031	−0.0371	−0.0046	−0.0422
	(−1.00)	**(−1.00)**	—	**(−1.00)**	(−0.80)	**(−1.00)**	(−0.60)	**(−1.00)**	(−0.60)	**(−1.00)**
HF$_{30}^{(1)}$	−0.0060	−0.0180	0.0213	—	0.0169	−0.0240	0.0201	−0.0139	0.0186	−0.0190
	(−0.40)	(−0.80)	**(1.00)**	—	(0.80)	**(−1.00)**	**(1.00)**	**(−1.00)**	**(1.00)**	**(−1.00)**
HF$_5$	−0.0205	−0.0326	0.0067	−0.0145	—	−0.0409	0.0032	−0.0308	0.0017	−0.0359
	(−0.80)	**(−1.00)**	(0.80)	(−0.80)	—	**(−1.00)**	(0.40)	**(−1.00)**	(0.00)	**(−1.00)**
HF$_5^{(10)}$	0.0053	−0.0067	0.0326	0.0113	0.0259	—	0.0440	0.0100	0.0425	0.0050
	(0.60)	(−0.40)	**(1.00)**	**(1.00)**	**(1.00)**	—	**(1.00)**	(0.60)	**(1.00)**	(0.40)
HF$_{75}$	−0.0276	−0.0396	−0.0003	−0.0216	−0.0071	−0.0329	—	−0.0340	−0.0015	−0.0391
	(−1.00)	**(−1.00)**	(0.20)	**(−1.00)**	(−0.20)	**(−1.00)**	—	**(−1.00)**	(−0.20)	**(−1.00)**
HF$_{75}^{(1)}$	0.0036	−0.0084	0.0309	0.0096	0.0241	−0.0017	0.0312	—	0.0325	−0.0051
	(0.40)	(−0.40)	**(1.00)**	**(1.00)**	**(1.00)**	(−0.20)	**(1.00)**	—	**(1.00)**	(0.00)
HF$_{130}$	−0.0262	−0.0382	0.0011	−0.0202	−0.0056	−0.0315	0.0015	−0.0297	—	−0.0376
	(−1.00)	**(−1.00)**	(0.20)	**(−1.00)**	(−0.20)	**(−1.00)**	(0.00)	**(−1.00)**	—	**(−1.00)**
HF$_{130}^{(1)}$	0.0021	−0.0099	0.0294	0.0081	0.0226	−0.0032	0.0297	−0.0015	0.0283	—
	(0.60)	(−0.40)	**(1.00)**	(0.40)	**(1.00)**	(−0.20)	**(1.00)**	(0.00)	**(1.00)**	—

The Precision of High-Frequency Betas

$$\beta^{\mathrm{R}}_{t,T} = a + b\beta_t + \epsilon_T,$$

Table 9. Univariate and Encompassing Regressions for Realized Beta

					Panel A: Univariate regressions							
	a	sig	b_{HIST}	sig	$b_{\text{HIST}^{(1)}}$	sig	$b_{\text{HF}_{30}}$	sig	$b_{\text{HF}_{30}^{(1)}}$	sig	R^2_{adj}	sig(Wald$_1$)
HIST	0.19	(0.20)	0.77	(0.20)	—	—	—	—	—	—	0.64	(1.00)
HIST$^{(1)}$	0.24	(0.80)	—	—	0.70	(1.00)	—	—	—	—	0.62	(1.00)
HF$_{30}$	0.13	(0.40)	—	—	—	—	0.88	(0.40)	—	—	0.73	(0.40)
HF$_{30}^{(1)}$	0.11	(0.20)	—	—	—	—	—	—	0.88	(0.20)	0.65	(0.80)

					Panel B: Encompassing regressions							
	a	sig	b_{HIST}	sig	$b_{\text{HIST}^{(1)}}$	sig	$b_{\text{HF}_{30}}$	sig	$b_{\text{HF}_{30}^{(1)}}$	sig	R^2_{adj}	sig(Wald$_1$)	sig(Wald$_2$)
HIST & HIST$^{(1)}$	0.20	(0.40)	0.38	(0.40)	0.37	(0.40)	—	—	—	—	0.65	(0.40)	(0.40)
HIST & HF$_{30}$	0.13	(0.40)	−0.01	(0.00)	—	—	0.88	(0.80)	—	—	0.73	(0.00)	(0.80)
HIST & HF$_{30}^{(1)}$	0.11	(0.20)	0.37	(0.80)	—	—	—	—	0.50	(0.80)	0.69	(0.80)	(0.80)
HIST$^{(1)}$ & HF$_{30}$	0.14	(0.40)	—	—	0.08	(0.00)	0.78	(0.80)	—	—	0.73	(0.00)	(0.80)
HIST$^{(1)}$ & HF$_{30}^{(1)}$	0.13	(0.20)	—	—	0.33	(0.60)	—	—	0.52	(0.80)	0.69	(0.60)	(0.80)
HF$_{30}$ & HF$_{30}^{(1)}$	0.15	(0.40)	—	—	—	—	0.98	(1.00)	−0.11	(0.00)	0.73	(1.00)	(0.00)

The Precision of High-Frequency Betas

$$\beta^{\mathrm{R}}_{j,t,T} = a_T + b_T \beta_{j,t} + \epsilon_{j,t,T},$$

Table 10. Cross-Sectional Predictability

	HIST	HIST$^{(1)}$	HF$_{30}$	HF$_{30}^{(1)}$
		Panel A: Rank correlations		
ρ	0.6937***	0.6308***	0.8034***	0.7700***
(p-value)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
		Panel B: Cross-sectional regressions		
Const.	−0.0871***	−0.0264	0.0411***	−0.0602***
(s.e.)	(0.0221)	(0.0251)	(0.0125)	(0.0161)
Slope	0.9547***	0.8352***	0.9670***	0.9662***
(s.e.)	(0.1654)	(0.1377)	(0.1106)	(0.1371)
R^2_{adj}	0.4903	0.3900	0.6613	0.5984

Conclusion

本文基于日数据和高频数据，用betas检验了条件CAPM。而对于日常数据，该模型无法解释规模异常和三个主要异常的六个组成部分中的三个，而条件CAPM可以解释规模、价值和动量异常，以及使用高频数据时六个组成部分投资组合中的五个。我们发现，高频贝塔的优越经济价值与未来实现贝塔的优越可预测性有关。从高频数据估计的贝塔值比从日收益数据估计的贝塔值要准确得多。我们发现，在Mincer–Zarnowitz回归中，从高频数据估计的Beta也优于基于每日数据的Beta，并且具有更好的横截面可预测性。