INSTITUTE
OF ECONOMIC STUDIES
Faculty of Social Sciences
Charles University

ON EMPIRICAL CHALLENGES IN FORECASTING MARKET BETAS IN CRYPTO MARKETS

Jan Šíla
Michael Mark
Ladislav Krištoufek

IES Working Paper 19/2022

Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00 Praha 1

E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Bibliographic information:
Šíla J., Mark M. and Krištoufek L. (2022): "On Empirical Challenges in Forecasting Market Betas in Crypto Markets" IES Working Papers 19/2022. IES FSV. Charles University.

This paper can be downloaded at: http://ies.fsv.cuni.cz

On Empirical Challenges in Forecasting Market Betas in Crypto Markets

Jan Šíla[a,c]
Michael Mark[b]
Ladislav Krištoufek[a]

[a]Institute of Economic Studies, Faculty of Social Sciences, Charles University & The Czech Academy of Sciences, Institute of Information Theory and Automation, Prague, Czech Republic.
[b]Chair of Operations, Economics and Strategy, Ecole Polytechnique Federale de Lausanne, Station 5, CH-1015 Lausanne, Switzerland.
[c]Email: jan.sila@fsv.cuni.cz

July 2022

Abstract:
This paper investigates the predictability of market betas for crypto assets. The market beta is the optimal weight of a short position in a simple two-asset portfolio hedging the market risk. Investors are therefore keen to forecast the market beta accurately. Estimating the market beta is a fundamental financial problem and we document pervasive empirical issues that arise in the emerging market of crypto assets. Although recent empirical results about US stocks suggest predictability of the future realized betas about 55%, predictability for the universe of crypto assets is at most 20%. Our results suggest that the crypto market betas are highly sensitive not only to the beta estimation method but also to the selection of the market index. Thus we also contribute to the discussion on the appropriate market representation.

JEL: C21,C53,C58,G12
Keywords: Asset pricing, CAPM, Market Beta, Cryptocurrency

Acknowledgements: This work was supported by the Cooperatio Program at Charles University, research area Economics. Support from the Czech Science Foundation (project 20-17295S) is also highly appreciated.

1. Introduction

Much of the financial research has been devoted to the market beta, the well-known building block of the Capital Asset Pricing Model (CAPM) (Sharpe 1964, Lintner 1965, Mossin 1966). Although the CAPM itself has many discussed and documented flaws (Jensen et al. 1972, Roll 1977, Fama and French 2004), the market beta is still important for investors hedging their market risk. An investor can lower the risk of their investment by taking a short position in a market index instrument. The variance of such a portfolio is minimized if the market short position is proportional to the asset's OLS market beta. Since investors wish to hedge the future returns of their portfolio, accurately predicting the out-of-sample betas is of great interest. Here, we contribute to the literature by investigating the predictability of market betas in crypto assets[1]. We find that the predictability of the future yearly market betas in the universe of crypto assets is generally much lower compared to recent stock results of Welch (2019, 2021). While the historical stock betas explain between 51% and 55% of the future betas, in the crypto universe of a similar size, it is less than 20%. Moreover, we consider a total of four market index implementations, and for three of them, the historical betas explain less than 10%.

The univariate market model estimating the beta depends only on the market return, thus the market representation choice is crucial. Even though the US stocks-related financial literature typically uses the CRSP market index (Bali et al. 2016), the emerging asset pricing studies of crypto-assets do not have such a consensus. Mishra and O'Brien (2005) and Bruner et al. (2008) discuss the consequences of the choice between global and home country market index on stocks and find that it is consequential for emerging markets, while it is not for developed countries. Our results suggest that the market index selection in crypto strongly determines the predictability of the future betas, thus resembling emerging markets. While forecasting the future market betas and the market risk factor loadings, we document pervasive empirical issues which are likely to appear with other factors in the future studies of conditional models.

Our results also show a significant underestimation of the "stability coefficient" due to the classical errors-in-variables (CEV) bias which substantially underestimates the optimal hedging ratio. While Welch (2021) reports a negligible 2 % CEV bias for stocks, it is between 16% and 39% for crypto. Nevertheless, we observe structural differences between the applied market indices and show that the S&P Cryptocurrency Broad Digital Market index delivers the best results.

[1] We stick to the more general "crypto assets" or "crypto" as a shorthand notation as opposed to maybe more popular but less precise "cryptocurrencies" as the crypto markets are formed of coins, various tokens, stablecoins, wrapped variants of coins (tokenized coins) on different blockchains, and various DeFi (decentralized finance) protocols.

1.1. State of the art

It is well-documented that the market betas are not stable over time; they are time-varying and non-stationary (Blume 1971, Bos and Newbold 1984, Groenewold and Fraser 1999, Lewellen and Nagel 2006, Adrian and Franzoni 2009, Engle 2016). Yet the assumption of time-invariant OLS betas is implicit in the multitude of asset pricing tests, where the risk factor loadings are estimated as a constant parameter of a market model regression. This assumption is also commonly used in practice when the market beta estimated with historical data is used to make inferences for a future period. Strikingly, the ample evidence of instability has had only little influence on how the betas are estimated, with the notable exception of Kelly et al. (2019). Brenner and Smidt (1977), Vasicek (1973) or Dimson (1979) considering alternative approaches to OLS. However, the topic has not received much attention since the work of Shanken (1992). Welch (2019, 2021) produces recent empirical results for the US stock market when he proposes a set of non-OLS methods to estimate future OLS market betas. These studies show that there is a certain degree of predictability in the CRSP stock index yearly betas. Welch (2021) reports that yearly historical betas explain about 55% of the variation in the future yearly betas and that slope-winsorizing of returns produces more reliable and consistent estimates of the OLS market beta compared to the naive OLS estimate. Even though we also find that the other methods are superior to naive OLS estimates, the predictability is generally much lower. Another issue is the CEV bias, which underestimates the hedging ratio. In the stocks study of Welch (2021), the CEV bias is estimated to be around 2%, but it is more than 10 times as large in crypto.

The CEV problem is a known issue in theoretical econometrics (Wooldridge 2015) as well as in financial applications (Blume 1970, Shanken 1992). When one uses a beta estimate to predict the OLS beta in the future period, the inherent error in measuring the independent variable brings attenuation bias to the predictive regression. In turn it consistently underestimates the market beta prediction. Several remedies to CEV were proposed, most recently Jegadeesh et al. (2019) introduced an instrumental variable approach that allows using individual assets while alleviating the CEV bias. At this stage, we do not attempt to find the best approach to control for the CEV bias but rather show the issues it brings about in a common asset pricing setup. Historically, Blume (1970) proposed to use portfolios as test assets to reduce CEV in the estimated betas for asset pricing tests. Ang et al. (2020) discuss the trade-off between using portfolios and individual assets when estimating risk premia in the cross-section. We analyze a panel of individual assets results as well as portfolios. Our results show that indeed the crypto portfolios, even when formed on noisy betas, suffer less from the CEV bias, yet still the general predictability of portfolios does not match that of individual stocks.

1.2. State of crypto research

We add to the emerging literature of asset pricing in crypto. The crypto market has grown considerably in size over the recent years, yet remains relatively unexplored. The capitalization of the market in 2021 matched the size of some of the largest economies in the world[2] and several prominent publicly listed companies are directly exposed to crypto as well[3]. More importantly, El Salvador made Bitcoin a legal tender in 2021, particularly for its ability to cheaply transfer remittances from abroad to its citizens. Härdle et al. (2020) offer an overview of the current state of knowledge and future research applications as we still understand very little of this nascent and developing market. There is a large body of literature discussing the place of Bitcoin, and crypto assets in general, as a speculative asset class (Baur et al. 2018), its perceived role as a safe haven (Shahzad et al. 2019, Urquhart and Zhang 2019), efficiency (Kristoufek 2018, Urquhart 2016) or market endogenous dynamics (Mark et al. 2020). A solid strand of literature investigates time-varying volatility of crypto assets (Katsiampa 2017, Corbet et al. 2019, Conrad et al. 2018, Dyhrberg 2016, Berentsen and Schär 2019). Crypto assets are also known for their high volatility (Dutta and Bouri 2022, Zhang and Li 2020), hence outliers are set to play an important role in the market beta estimation procedure. Grané and Veiga (2010), Carnero et al. (2012) and Chang et al. (1988) document how outliers distort financial models in general. The high occurrence of return outliers, along with a quite volatile market index itself, raises challenges in estimating the market betas. The current crypto literature lacks a description of these issues for asset pricing. This also motivates considering other methods than OLS for estimating the market betas. We put these limitations together with the general problem of estimating market beta in a coherent framework. We compare methods presented in Welch (2021) that are robust to outliers and consider four different market indices from academia and industry. Our results demonstrate very limited stability of the market betas.

This is also the first study of the empirical differences in crypto market indices. The market index is a critical component in the vast majority of contemporary asset pricing models. However, there is an acute lack of consensus on which of the several crypto market indices is the best representative (Trimborn and Härdle 2018, Häusler and Xia 2022). As it turns out, differences in the number of constituents, requirements on their liquidity and in the weighting scheme lead to dramatically

[2] Total market capitalization of crypto in 2021 was between 1.5 and 3 trillion USD, which compares to the 2020 GDP of OECD countries – in trillions USD – Czech republic 0.46, Canada 1.75, Turkey 2.3, Italy 2.5, United Kingdom 3.14

[3] Microstrategy (MSTR) has large Bitcoin reserves, Coinbase (COIN) is one of the most prominent crypto exchanges, Bitfarms (BITF) specializes in hardware mining of the digital currencies. Notably, Tesla (TSLA), an S&P 500 constituent, has started accepting Bitcoin as a form of payment for their electric vehicles, although it stopped soon after.

different conclusions. Our results thus point to potential and undocumented issues with estimating the cross-sectional conditional factor models, which is a common topic of financial research (Fama and MacBeth 1973, Fama and French 1993, Gebhardt et al. 2005, Harvey et al. 2016).

If risk factor betas are highly variable and noisy, then constructing portfolios based on their ranks implies unreliable dispersion in returns as well as post-rank loadings. The current crypto asset pricing studies focus on explaining the variation in the cross-section of crypto returns (Liu et al. 2022, Shen et al. 2020, Bianchi and Babiak 2021, Zhang and Li 2020) using crypto characteristics and we are not aware of any study considering factor loadings. Only Liu and Tsyvinski (2021) use factor loadings, but only to explain monthly crypto market returns, not individual asset's expected returns and they do not consider crypto specific financial factors. Asset pricing in crypto is thus missing an examination of a very relevant financial problem. Moreover, asset pricing tests based on beta ranks commonly use OLS betas and shorter estimation periods. Ang et al. (2006) state that using a monthly historical window to predict future factors loadings over a month is a natural compromise. We do not find stability in market betas for any estimation method or index for quarterly and monthly periods. Shorter estimation horizon further deteriorates the already poor forecast ability across all of the specifications in this study.

1.3. Outline

The remainder of this paper is organized as follows. In Section 2, we describe the methodology used for estimating market betas. Section 3 provides an empirical description of our dataset as well as a discussion on selecting the cryptocurrency market index. In Section 4, we present the forward-looking performance of the various beta estimators. Finally, Section 5 concludes.

2. Methodology

We investigate and compare the performance of various estimators in forecasting the future OLS market betas. An investor can only use historically available information in estimating a future and unobservable market beta. The selection of estimators is not limited to OLS, yet the OLS as the best linear unbiased estimator is the natural target. Welch (2019, 2021) discusses several other methods and find them to be superior to OLS in the predictive performance of the future OLS betas. In our methodological framework, we select the three best-performing ones from Welch (2021) as alternatives to the OLS betas. Namely, we consider plain slope winsorized betas, age decaying slope winsorized betas, and Vasicek (1973) betas that shrink beta estimates towards a Bayesian prior.

Winsorization and truncation are common techniques to handle outliers in empirical asset pricing (Bali et al. 2016). Welch (2019) introduces slope winsorization which limits the security's return based on the contemporaneous market return. The implementation is very simple and only requires a single pass through the data to process the in-sample security returns. Throughout the paper, we denote plain slope winsorized betas as β^{SW}, while the OLS estimate is β^{OLS}. In general, the time t winsorized excess return $r_{i,t}^{\mathrm{SW}}$ of a security i is required to belong to a bounded interval

$$r_{i,t}^{\mathrm{SW}} \in \big(r_{m,t}(1-\Delta), r_{m,t}(1+\Delta)\big), \tag{1}$$

where $r_{m,t}$ is the market return at time t and Δ is the winsorization parameter. Following Welch (2021), we impose $\Delta \triangleq 3$. Therefore, the allowed return range is limited to $r_{i,t}^{\mathrm{SW}} \in (-2 \cdot r_{m,t}, 4 \cdot r_{m,t})$. The estimator of the slope winsorized beta β^{SW} is then defined as

$$\hat{\beta}_i^{\mathrm{SW}} = \frac{\mathrm{Cov}(r_{i,t}^{\mathrm{SW}}, r_{m,t})}{\mathrm{Var}(r_{m,t})}. \tag{2}$$

Similarly, the age-decayed slope winsorization beta β^{SWA} uses the same winsorization range but is estimated with weighted-least-squares (WLS) regression that smooths the time decay. We use β^{SWA} for the yearly beta estimates with the same parameters as Welch (2021), such that WLS weights decay is roughly 0.55% per day. Thus the half-live of observations is 125 days and the weight of yesterday's return is about twice as high as four months ago and about 8 times the weight of the return a year ago.

Lastly, we implement the Bayesian shrinkage estimator introduced by Vasicek (1973). The Vasicek adjustment shifts the OLS beta towards a cross-sectional expectation such that the magnitude of the shift is increasing with the standard error of the OLS estimate. That is, for precise estimates Vasicek shrinkage gives more weight to the OLS beta while for noisy estimates it pushes the beta towards a cross-sectional expectation. It requires first calculating a panel of individual OLS market betas and recording the standard error of the estimations. We calculate a cross-sectional mean of the betas $\overline{\hat{\beta}^{\mathrm{OLS}}}$ and standard deviations $\overline{\sigma_t^2}$. Then the Vasicek estimate $\hat{\beta}^{\mathrm{VCK}}$ follows;

$$\hat{\beta}_{i,t}^{\mathrm{VCK}} = w_i \cdot \hat{\beta}_i^{\mathrm{OLS}} + (1 - w_i) \cdot \overline{\hat{\beta}_t^{\mathrm{OLS}}} \tag{3}$$

where $w_i = \frac{\overline{\sigma_t^2}}{\sigma_i^2 + \overline{\sigma_t^2}}$ is the weight of the shrinkage. Empirical evidence mentioned in Welch (2019) suggests that Vasicek's betas perform well with time-varying underlying betas and in presence of outliers, which is our case.

Investors are keen to be able to accurately forecast future realized beta values to optimize hedging the market risk. Hence, we forecast $\hat{\beta}_{i,t+1}^{\mathrm{OLS}}$ with $\hat{\beta}_{i,t}^{(\cdot)}$, where $\hat{\beta}_{i,t}^{(\cdot)}$ stands for $\hat{\beta}^{\mathrm{OLS}}$, $\hat{\beta}^{\mathrm{SW}}$, $\hat{\beta}^{\mathrm{SWA}}$ or $\hat{\beta}^{\mathrm{VCK}}$. We judge the quality of the predictions with the R^2 from a pooled panel data regression

$$\hat{\beta}_{i,t+1}^{\mathrm{OLS}} = \gamma_0 + \gamma_{\hat{\beta}} \hat{\beta}_{i,t}^{(\cdot)} + \epsilon_{i,t+1}, \tag{4}$$

and with the Root Mean Square Error (RMSE) of the future beta predictor over all successive periods $\tau \in \{1,..,T\}$ and assets $i \in \{1,...,n\}$, i.e.,

$$\text{RMSE} = \sqrt{\sum_i \sum_\tau \frac{(\hat{\beta}^{\text{OLS}}_{i,\tau+1} - \hat{\beta}^{(\cdot)}_{i,\tau})^2}{nT}}. \tag{5}$$

Note that R^2 is a measure of beta predictability and we interpret $\gamma_{\hat{\beta}}$ as a "stability coefficient". Indeed, if the betas were perfectly persistent, $\gamma_{\hat{\beta}}$ would be close to unity and on average, investors would just keep the historical hedging ratios. Since the explanatory variable in the pooled regression in Equation (4) is an estimate with an inherent measurement error, we face the classical error-in-variables problem (CEV), or rather the attenuation bias (Wooldridge 2015), which underestimates $\gamma_{\hat{\beta}}$. Welch (2021) notes that while RMSE is affected by this bias, R^2 is not and hence it is a more suitable measure when comparing the results of predictability. We also include the Mean Absolute Error metric (MAE) in the results, primarily to discern different sources of prediction errors.

To illustrate the CEV issue, let us consider an explanatory variable x measured with an error u, and let $\tilde{x} = x + u$ be the regressor in a general univariate OLS regression. Then the model follows $y = \beta(\tilde{x} - u) + \epsilon = \beta\tilde{x} + (\epsilon - \beta u)$, where the measurement error becomes a part of the error term creating an endogeneous bias. Then the OLS estimator can be easily derived (Wooldridge 2015) to follow

$$\hat{\beta} \underset{p}{\rightarrow} \beta \frac{\delta_x^2}{\delta_x^2 + \delta_u^2} = \beta\lambda, \tag{6}$$

where λ is the attenuation factor. Following Welch (2021), we calculate δ_x as the average dispersion of the estimated betas in the cross-section and δ_u as the average standard error of estimating the beta for the estimator, market index, and holding period, respectively. We report the value $1 - \lambda$ as the CEV bias, which is about 2% in the stock sample published in Welch (2019, 2021).

3. Data

As the data structures, possibilities but also issues and pitfalls are quite different from the standard financial instruments, we attribute much attention to proper and precise description to the specific empirical challenges that crypto data pose in research. We downloaded the full history of all coins and tokens (over 18,000 assets) that were ever published on CoinMarketCap.com (CMC), a standard data source in the crypto literature (Liu and Tsyvinski 2021, Liu et al. 2022, Zhang and Li 2020, Kristoufek and Vosvrda 2019, Kristoufek 2021, Kukacka and Kristoufek 2020, Shen et al. 2020). Due to the availability of the market indices, we run the analysis from August 2014 to December 2021. After downloading the data, we exclude IDs that are re-used in more projects, which we found in several cases, and consider this an internal error of the data provider. We eliminate wrapped coins, stablecoins, and leveraged tokens depending on tags provided by CMC and full-text search of keywords in the names, as the tags are not entirely consistent.

3.1. Data properties

One of the challenges is that typically, crypto assets have returns larger than 5 or even 7 sigmas of their respective distributions. Commonly, the raw dataset contains assets with daily returns in the hundreds of percents with volumes of less than 10 USD. Such extreme returns are removed from the dataset if the corresponding volume for that return is less than 1 milion USD as this is likely to be a result of the so-called "pump and dump" or other market manipulation (Kamps and Kleinberg 2018, Li et al. 2021) occurring in illiquid assets. Also, we eliminate individual observations with infinite returns or returns corresponding to zero prices. Finally, we interpolate at most one missing price observation, otherwise, the asset is removed from the sample. Surprisingly, those empirical properties of the crypto data are not mentioned in the current literature (Liu and Tsyvinski 2021, Liu et al. 2022, Shen et al. 2020, Zhang and Li 2020, Bianchi and Babiak 2021) while the studies use the same data provider.

After the initial processing, there are 4,883 coins and tokens in the considered universe of crypto assets. Of those, 3,483 assets (71%) ever record a daily return of more than 100% and 1,214 assets (24%) even a daily return larger than 500%. Each asset in our universe has at least one absolute daily return greater than three sigmas and 3,870 assets (79%) greater than seven sigmas. For example, Bitcoin's standard deviation of all daily returns is 3.87% with the lowest return -37% and the highest daily gain of 25%. The return characteristics are also described in (Liu and Tsyvinski 2021) and we compare them with the market indices in Table 2. We expect those properties to present challenges, particularly for the β^{OLS} estimation.

Winsorization of the returns is an option to handle such returns. Table 1 shows what proportion of returns are winsorized, respectively slope-winsorized as in Welch (2021) for each of the indices introduced in Subsection 3.2. From now on, we consider asset and market returns to be net of the daily risk free rate downloaded from Ken French's data library. On average an asset has between 31% and 37% of the daily returns outside of the slope-winsorization market return band. Welch (2021) does not report what proportion of the stock returns is winsorized, but we can reasonably expect that it will be much lower. The largest coins have the lowest number of winsorized returns as, by design, they influence the market indices the most and thus correlate with them the most.

Nevertheless, the descriptive statistics illustrate that even simple empirical results differ dramatically across the market indices. For example, Bitcoin has 5.5% of its daily returns in the sample winsorized when compared to BDMI, in which it is not actually a constituent, while this number almost doubles when considering CCMIX. Such differences already hint that the forecasting performances can vary considerably across the applied indices.

	CCMIX	CRIX	SCI	BDMI
Min	9.5%	6.8%	5.6%	5.5%
Mean	37.1%	37.2%	35.7%	31.4%
Median	36.3%	36.6%	34.7%	30.5%
Max	64.9%	63.9%	63.0%	61.5%

Table 1 Aggregate statistics of winsorized returns. The table shows the distribution of the proportion of winsorized returns depending on the market index used. An average asset has about 37 % of daily returns winsorized.

3.2. Selecting the market index

Standard asset pricing literature typically uses the CRSP value-weighted market portfolio from Ken French's data library as the market index. Yet the limited literature on crypto asset pricing has no such consensus despite its critical role in the majority of financial research. Trimborn and Härdle (2018) summarize the main challenges in constructing a relevant and descriptive market index for cryptocurrencies. Unlike the stock market, crypto is traded every day around the clock and across hundreds of exchanges with varying liquidity. Trimborn and Härdle (2018) introduce a methodology resulting in the CRIX index which accounts for those specifics as well as typically the fast rise and demise of crypto projects. Following the CRSP approach, the CRIX is dynamic in the number of constituents. CRIX only adds an index member slot, if it improves the AIC criterion of the index, thus it penalizes high number of index members. The number of constituents is evaluated every quarter, but the individual constituents are selected and rebalanced monthly. This allows the index to adapt to rapidly changing market conditions.

Second index considered is the Broad Digital Market Index (BDMI)[4] produced by S&P Digital market division. It rebalances on the last calendar days of February, May, August, and November. Each asset needs to have a market capitalization of at least 10 million USD and have a three-month median daily value traded (MDVT) of at least 100 thousand USD. Then the index is a value-weighted portfolio of the eligible constituents. Unlike all the other indices, it excludes Bitcoin and Ethereum to account more for the behavior of the broad market's alternative assets which are much lower in capitalization. Since the online published time series does not include weekends, we recalculated the index following the methodology so that it corresponds to our dataset.

The Crescent Crypto Market Index (CCMIX)[5] originates in industry, but it has been used in academic works (Shah et al. 2021, Ramos et al. 2021) as well. Despite being discontinued at the end of 2021, we include it for a more complete historical comparison. It is an example of a directly

[4] Bloomgerg ticker "SPCBDM", Reuters ".SPCBDM"

[5] Bloomberg ticker "CCMIX", Reuters ".CCMIX"

investable index managed by the Crescent Crypto Manager LLC that includes at most the twenty largest value-weighted coins and is rebalanced monthly.

Lastly, we consider a pure value-weighted portfolio of all assets in the universe. Although Trimborn and Härdle (2018) published a market index suitable for the dynamic nascent crypto market, recent literature (Zhang and Li 2020, Bianchi and Babiak 2021, Liu and Tsyvinski 2021, Liu et al. 2022) rather considers a value-weighted portfolio of all the assets in the universe, with minimal selection rules, perhaps due to the mentioned lack of consensus. More specifically, Zhang and Li (2020) consider cryptocurrencies with at least 2 years of trading record, Liu and Tsyvinski (2021) and Liu et al. (2022) assets with capitalization above 1 million USD, and Bianchi and Babiak (2021) assets with capitalization over 100 million USD which are missing no more than 25% of observations in the period. The articles do not mention if the capitalization levels are required for the whole period, or if the asset needs to attain it at least once over the observed sample. Also no details are provided with respect to rebalancing, so we assume that weights are assigned daily, which makes it, unlike the other indices, not directly investible due to the number and costs of daily rebalancing transactions. Surprisingly, other papers pose no restrictions with respect to the volume traded, which we find quite important for judging the crypto's relevance and eliminating some very suspiciously behaving assets. Nevertheless, for our construction, we require the assets to have a daily average market capitalization of at least 1 million USD and a daily median volume of at least 10 000 USD to eliminate assets with a very high percentage of large outliers. We label this index in the subsequent analysis as Simple Crypto Index (SCI).

Even though the indices attempt to track the same market, they exhibit very different empirical characteristics. Table 2 documents basic statistics of the market indices and two major cryptocurrencies, Bitcoin and Ethereum, from August 2014 to December 2021. There are notable differences between indices particularly in the annual returns. For example, SCI gains 10,000 % in the bull market of 2017, as a very broad index which contains lots of booming assets gaining hundreds of percent weekly, led mainly by the rise of Ethereum. In the same period, CRIX and CCMIX are much more conservative with return profiles more similar to Bitcoin in the daily and compounded returns. One of the reasons is that SCI is rebalanced daily with no assumed transaction costs, which is a rather hypothetical portfolio than a practically investable one. However, the mentioned papers do not consider the practical feasibility of their market index. This approach suffers from asymmetry in returns, as while an asset's downside is naturally limited, many crypto assets attain returns so large that they contribute substantially to the index return despite their very small weight. The distributions of daily returns in the given year have consistently higher mean and median returns compared to the stock market as well as respective standard deviations of those returns. The mean returns tend to be further from zero than the median suggesting heavy tails in

Index		2014	2015	2016	2017	2018	2019	2020	2021
BDMI	Mean	0.31	-0.17	0.12	1.33	-0.45	-0.05	0.27	0.62
	Median	0.07	-0.23	0.02	0.79	-0.29	0.10	0.59	1.04
	Standard deviation	4.50	3.85	2.15	5.91	5.81	3.62	4.51	5.45
	Total Return	37.40	-59.36	40.60	6,645.89	-89.71	-33.78	8.31	448.67
CCMIX	Mean				0.94	-0.32	0.14	0.44	0.35
	Median				1.15	-0.08	0.01	0.46	0.40
	Standard deviation				4.61	4.81	3.77	3.57	4.47
	Total Return				1,977.54	-79.85	28.87	295.62	135.28
CRIX	Mean	-0.27	0.10	0.26	1.02	-0.35	0.17	0.51	0.52
	Median	-0.40	0.13	0.20	0.99	-0.04	0.15	0.39	0.32
	Standard deviation	2.99	3.46	2.45	4.65	4.68	3.56	3.76	4.96
	Total Return	-37.85	16.06	131.73	2,654.16	-87.70	47.23	392.42	337.18
SCI	Mean	-0.20	0.16	0.40	1.39	-0.17	0.26	0.62	0.64
	Median	-0.35	0.18	0.25	1.44	-0.01	0.19	0.54	0.78
	Standard deviation	3.05	3.47	2.41	4.76	4.94	3.51	4.08	4.63
	Total Return	-30.94	44.84	288.18	10,045.42	-66.14	102.94	585.73	586.37
Major Crypto Assets									
Bitcoin	Mean	-0.35	0.15	0.25	0.86	-0.27	0.24	0.46	0.22
	Median	-0.41	0.12	0.18	0.89	0.08	0.12	0.27	0.13
	Standard deviation	3.10	3.60	2.51	4.99	4.24	3.56	3.77	4.21
	Total Return	-45.38	34.47	123.83	1,368.90	-73.56	92.20	303.16	59.69
Ethereum	Mean		0.56	0.82	1.51	-0.32	0.08	0.61	0.60
	Median		-0.01	-0.19	0.39	-0.26	-0.08	0.53	0.59
	Standard deviation		9.62	6.92	7.30	5.60	4.11	4.94	5.60
	Total Return		23.92	753.64	9,395.84	-82.38	-2.82	469.25	399.13

Table 2 The table shows mean, median and standard deviation of daily returns and total cumulative return per year of four market indices and two largest cryptocurrencies from August 2014 to the end of 2021. All values are in percentages, including the total returns for consistency. All the indices start on August 1, 2014 and run to the end of 2021. CCMIX runs from January 1, 2017 to December 12, 2021.

the distributions; extremely so in the case of Ethereum and BDMI in 2017, where the mean return is 1.5% daily compared to the median 0.4%, or rather 1.33% to 0.8% in the case of BDMI, which excludes Ethereum and Bitcoin, but rather tracks the so-called altcoins. This effect however wears out in the latest bull run of 2020 and 2021.

4. Results

We investigate the stability of market beta coefficients, or rather their predictability, using various methods and market indices from August 2014 to December 2021. Primarily, we are interested in predicting the yearly market betas, as one of the use cases of market beta is hedging against market risk. We also consider shorter horizons as they are more typical for asset pricing tests. Due to the short history of crypto, rather than calculating betas at the calendar year-end, we estimate

Index	Predictor	T	$1-\lambda$	γ_0	$\gamma_{\hat{\beta}}$	SD	$\gamma_{\hat{\beta}}/\lambda$	R^2	RMSE	MAE
BDMI	$\hat{\beta}^{\text{OLS}}$	12,495	0.388	0.451	0.423	0.001	0.659	0.163	0.269	0.195
	$\hat{\beta}^{\text{SW}}$		0.168	0.333	0.563	0.001	0.665	0.203	0.241	0.179
	$\hat{\beta}^{\text{SWA}}$		0.131	0.376	0.515	0.001	0.585	0.194	0.248	0.184
	$\hat{\beta}^{\text{VCK}}$		0.315	0.260	0.643	0.001	0.930	0.186	0.237	0.176
CCMIX	$\hat{\beta}^{\text{OLS}}$	10,879	0.413	0.622	0.207	0.001	0.305	0.044	0.313	0.234
	$\hat{\beta}^{\text{SW}}$		0.183	0.566	0.273	0.001	0.307	0.050	0.284	0.217
	$\hat{\beta}^{\text{SWA}}$		0.143	0.600	0.234	0.001	0.275	0.040	0.293	0.227
	$\hat{\beta}^{\text{VCK}}$		0.322	0.577	0.256	0.001	0.333	0.032	0.279	0.212
CRIX	$\hat{\beta}^{\text{OLS}}$	12,495	0.327	0.579	0.250	0.001	0.356	0.059	0.339	0.258
	$\hat{\beta}^{\text{SW}}$		0.123	0.509	0.329	0.001	0.368	0.071	0.310	0.242
	$\hat{\beta}^{\text{SWA}}$		0.114	0.540	0.300	0.001	0.347	0.064	0.312	0.243
	$\hat{\beta}^{\text{VCK}}$		0.245	0.498	0.345	0.001	0.453	0.056	0.298	0.231
SCI	$\hat{\beta}^{\text{OLS}}$	12,495	0.383	0.615	0.311	0.001	0.449	0.088	0.307	0.224
	$\hat{\beta}^{\text{SW}}$		0.157	0.546	0.388	0.001	0.438	0.096	0.283	0.210
	$\hat{\beta}^{\text{SWA}}$		0.126	0.558	0.380	0.001	0.429	0.101	0.286	0.211
	$\hat{\beta}^{\text{VCK}}$		0.308	0.523	0.418	0.001	0.558	0.080	0.275	0.203

Table 3 This table shows yearly overlapping beta predictions of the future OLS beta. T represents the number of observations in the panel, where CCMIX has a shorter time span. γ_0 is the intercept and $\gamma_{\hat{\beta}}$ the coefficient of interest from Equation (4) with respective standard error in column SD. The $1-\lambda$ is an estimate of the CEV bias underestimating the $\gamma_{\hat{\beta}}$, while $\gamma_{\hat{\beta}}/\lambda$ represents the rescaled coefficient. R^2 is the standard measure of fit of the pooled panel regression and RMSE and MAE are fit metrics introduced in Section 2.

the betas at every calendar month-end. Therefore, we consider a year to be 360 days, rather than a calendar year to have a consistent estimation window independent of the start. The primary goal of the current study is to examine how well the different estimation methods forecast future OLS betas. We forecast $\hat{\beta}^{\text{OLS}}_{i,t+1}$ with an estimator's $\hat{\beta}^{(\cdot)}_{i,t}$ realization, where $\hat{\beta}^{(\cdot)}_{i,t}$ stands for $\hat{\beta}^{\text{OLS}}$, $\hat{\beta}^{\text{SW}}$, $\hat{\beta}^{\text{SWA}}$ or $\hat{\beta}^{\text{VCK}}$.

4.1. Forecasting individual market betas

The main result is summarized in Table 3 where we show the results of pooled regression predicting one-year ahead beta. We use overlapping observation in the sense that we predict beta for Jan-Dec 2017 with daily returns from Jan-Dec 2016, then Feb 2017 - Jan 2018 betas with Feb 2016 - Jan 2017 data, and so on. Thus, we adjust the gamma standard errors with the appropriate Newey-West corrections. The variation in the forecasting metrics is very large across the indices and methods. With respect to the regression metrics in the pooled regression, the $\gamma_{\hat{\beta}}$ coefficient from Equation (4) is always highly significant and as expected positive. While Welch (2021) reports $\gamma_{\hat{\beta}}$ coefficients between 0.7 and 0.9 for the same methods, for cryptos, this coefficients is much lower. This is due to the CEV bias reported in $1-\lambda$ column, which underestimates the $\gamma_{\hat{\beta}}$. Hence, the column $\gamma_{\hat{\beta}}/\lambda$ represents the $\gamma_{\hat{\beta}}$ rescaled by the attenuation factor. Note that λ is again only an estimate of the CEV bias but even controlling for it in CCMIX, CRIX and SCI does not estimate the stability coefficient close to values reported in Welch (2021). This suggests there are also other factors driving the poor predictability of crypto market betas.

The R^2 of 20% for $\hat{\beta}^{\mathrm{SW}}$ in BDMI is the best forecasting performance, yet still very poor compared to the predictability of US stock betas in Welch (2019, 2021) which is between 58% and 60% for the 1,000 largest stocks and 51% and 55% for the 3,000 largest, respectively. The results show the superior performance of the winsorized methods and Vasicek's betas compared to OLS, identically to Welch (2019, 2021). Yet the R^2 measured across the market indices and estimators are notably lower than in the CRSP stock dataset, while the RMSEs are similar. Out of the four considered market indices, BDMI has consistently higher $\gamma_{\hat{\beta}}$ as well as R^2 higher almost an order of magnitude compared to the others, where the R^2 is abysmally low.

The R^2 measured in our case are much smaller compared to Welch (2021). While for instance his OLS betas explain 58% in the future OLS beta variation, in BDMI it is only 16%, respectively 4.4%, 5.9% and 8.8% in the other indices. The highest R^2 in Table 3 is $\hat{\beta}^{\mathrm{SW}}$ estimator in BDMI, which is 20.3%. Those R^2 results correspond to stocks' market betas predictions for 5 and 10 years into the future, not the subsequent period. Nevertheless, we observe that $\hat{\beta}^{\mathrm{SW}}$ always improves the forecasting over OLS, in the case of BDMI by 24%, for CCMIX by 13% and 20% for CRIX, 9% respectively for SCI. Unlike Welch (2021) who measures the attenuation bias $1-\lambda$ to be around 2%, we estimate much higher values for crypto. For $\hat{\beta}^{\mathrm{OLS}}$ betas, it is between 32% and 39% across the indices. While it is about half for the other methods, it is still significantly higher than in the Welch's similar study. Nevertheless, the poor performance of the forecasting for different indices cannot be attributed solely to the CEV problem as it is relatively similar within an estimation method. Neither does it seem to explain why SCI has higher consistency of market betas relative to the other indices. Neither the RMSE nor the MAE metrics present such a structural difference, although they are affected by the CEV bias. We hypothesize that it is due to individual market index properties, which makes the research into the comprehensive market representation of crypto urgent.

Overall, the time variation in the pooled betas is extreme. Figure 1 illustrates several prominent crypto-assets and their SCI betas, the index most used in the literature. Colors indicate in which calendar year the historical data for estimation end and we observe much different time dependent dynamics per asset. While the most important assets Bitcoin and Ethereum cluster around 1, XRP and Doge have very dispersed values and clearly do not exhibit any consistent linear relationship between the historical and future realized market betas throughout the dataset. Ethereum and XRP exhibit a distinct 2016-2017 cluster, corresponding to the then bull market. Doge shows a

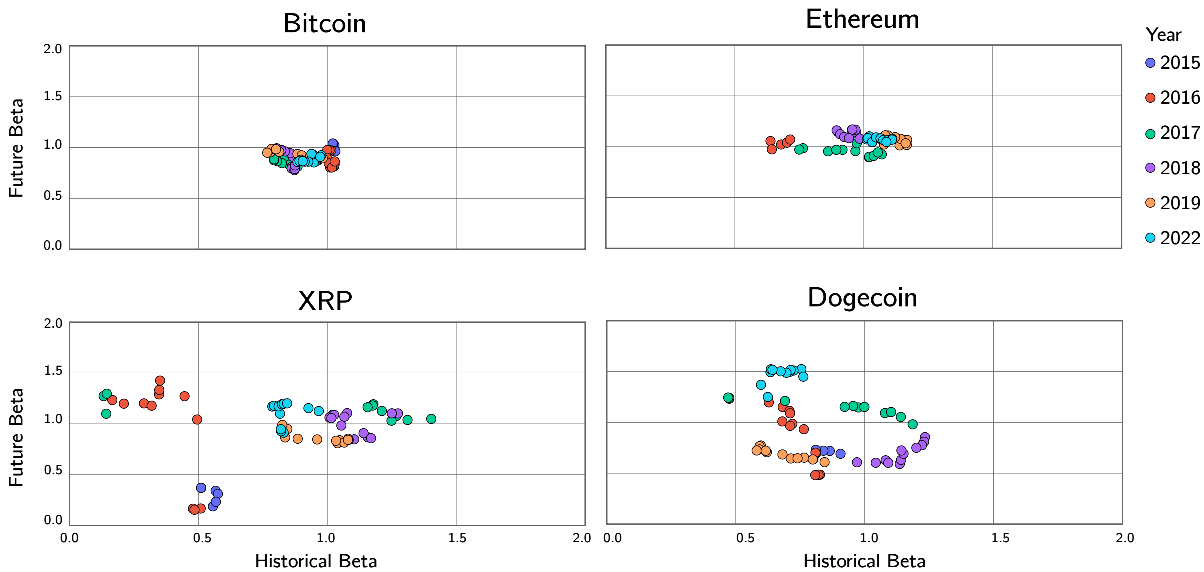

Figure 1 This figure show the subsample from the pooled beta dataset for four prominent cryptos. Colors indicate in which calendar year the historical data for estimation start. The estimated historical beta is on the x-axis and the future realized beta (to be predicted based on the historical one) is on the y-axis.

cluster of high 2021 betas in blue affected by Elon Musk's adoption of Doge as a personal favourite crypto asset in Spring 2021, increasing its price almost 30 times over a few months. XRP shows very volatile behavior, particularly in the first half of the sample. This points to the issue of quickly changing issues of the relative importance of assets discussed in (Trimborn and Härdle 2018). Although illustrative, the figure suggests that crypto assets have individually very different market beta dynamics and those pooled together yield the documented inaccuracies. Groenewold and Fraser (1999) report time trend in market betas as an important factor in their study of the Australian sector portfolios and they identify significant structural break in 1987. Crypto has experienced profound bear markets and explosive bull markets which was driven primarily by altcoins such as XRP and Dogecoin. A potential avenue of future research would be to explore this time dependency and dissect the crypto-asset universe by characteristics to investigate specific stability in the sub-groups.

4.2. Forecasting market betas of portfolios

Analyzing portfolios instead of individual assets is expected to alleviate the CEV problem (Blume 1970). Table 4 shows the same analysis on portfolios consisting of the same assets as previously. For this exercise, we disregard transaction costs, as we focus on portfolio market beta prediction rather than profitability. At the end of each month, we calculate historical yearly market betas for each market index and estimation method. The assets are sorted in quintile portfolios based on their market betas and then value-weighted and held for the same time period in the future. This is also basis of the typical asset pricing procedure of estimating risk premia in the cross-section based on risk factor loadings. Table 4 compares the predictive power of the portfolio's historical beta to the future portfolio's OLS beta.

As expected, using portfolios instead of individual assets brings a significant reduction in the estimated CEV bias, driven by lower dispersion in the pooled betas. Surprisingly, the $\gamma_{\hat{\beta}}$ are not much higher. Portfolios significantly improve the predictability in terms of R^2. All the $\gamma_{\hat{\beta}}$ coefficients are strongly statistically significant. For the portfolios, the $1-\lambda$ drops by an order of magnitude and ranges in units of percent, while R^2 increases, particularly in the poorly performing indices where it almost doubles. For example, $\hat{\beta}^{\text{SWA}}$ in CCMIX and CRIX reaches R^2 of 20% compared to 5%, respectively 6.4%. The stability of SCI market betas improves the least and $\hat{\beta}^{\text{VCK}}$ betas explain only 14.6% in the future variation of OLS betas. While for BDMI, the improved R^2 attains 50%. Nevertheless, even when using portfolios, our predictability of crypto market betas is worse than in Welch (2019, 2021) who uses individual stocks. As for the estimators, the OLS betas again do not perform well; they have consistently the highest CEV bias and the lowest R^2. Notably, the $\hat{\beta}^{\text{SW}}$ estimator actually performs worse than OLS in terms of R^2 with $\hat{\beta}^{\text{VCK}}$ or $\hat{\beta}^{\text{SWA}}$ having the highest explanatory power. The results in Table 3 and Table 4 show a persistent heterogeneity across the market indices and the superiority of non-OLS methods in predicting the future OLS market beta. Throughout the measurements, the BDMI index performs the best, compared to the plain value-weighted portfolio of SCI. The BDMI index puts restrictions on market capitalization and historical trading volume and appears to be the most suitable for further asset pricing research.

4.3. Forecasting shorter horizons

We also analyze horizons of 180 and 30 days. We consider the shorter horizon to describe the potential instability of the beta coefficients for portfolio sorts tests such as (Ang et al. 2006, Chang et al. 2013), where reasonable stability of the post-ranking factor loadings is required. Liu et al. (2022), Zhang and Li (2020), Shen et al. (2020) study the pricing of factors even in weekly returns for crypto, however, we omit weekly beta estimates in our analysis. Again, we predict the following out of sample period and compare market indices and different estimation methods trying to forecast the OLS beta. Due to much shorter estimation periods, we omitted the age decay winsorization method β^{SWA}. Table 5 summarizes the results.

The predictive performance degrades significantly with shorter horizons. The same observation is noted in Welch (2019), but no numerical results are published. Consistent with the yearly betas in

Index	Predictor	T	$1-\lambda$	γ_0	$\gamma_{\hat{\beta}}$	SD	$\gamma_{\hat{\beta}}/\lambda$	R^2	RMSE	MAE
BDMI	$\hat{\beta}^{\text{OLS}}$	330	0.076	0.430	0.529	0.003	0.573	0.473	0.214	0.160
	$\hat{\beta}^{\text{SW}}$		0.064	0.385	0.586	0.004	0.626	0.461	0.196	0.149
	$\hat{\beta}^{\text{SWA}}$		0.052	0.379	0.591	0.007	0.624	0.448	0.207	0.151
	$\hat{\beta}^{\text{VCK}}$		0.068	0.407	0.556	0.003	0.597	0.507	0.176	0.135
CCMIX	$\hat{\beta}^{\text{OLS}}$	180	0.090	0.613	0.277	0.010	0.304	0.197	0.192	0.153
	$\hat{\beta}^{\text{SW}}$		0.068	0.637	0.252	0.011	0.270	0.164	0.183	0.146
	$\hat{\beta}^{\text{SWA}}$		0.060	0.630	0.264	0.007	0.281	0.221	0.175	0.143
	$\hat{\beta}^{\text{VCK}}$		0.066	0.626	0.276	0.010	0.296	0.211	0.167	0.134
CRIX	$\hat{\beta}^{\text{OLS}}$	325	0.088	0.599	0.343	0.009	0.376	0.185	0.245	0.189
	$\hat{\beta}^{\text{SW}}$		0.045	0.588	0.352	0.010	0.369	0.178	0.239	0.185
	$\hat{\beta}^{\text{SWA}}$		0.038	0.588	0.356	0.010	0.370	0.220	0.240	0.180
	$\hat{\beta}^{\text{VCK}}$		0.064	0.602	0.337	0.007	0.360	0.183	0.214	0.166
SCI	$\hat{\beta}^{\text{OLS}}$	335	0.096	0.695	0.266	0.010	0.294	0.122	0.253	0.186
	$\hat{\beta}^{\text{SW}}$		0.051	0.686	0.275	0.010	0.290	0.118	0.246	0.183
	$\hat{\beta}^{\text{SWA}}$		0.041	0.689	0.274	0.011	0.286	0.136	0.255	0.183
	$\hat{\beta}^{\text{VCK}}$		0.070	0.670	0.286	0.006	0.308	0.146	0.216	0.162

Table 4 Panel of quintile portfolio betas. This table shows the same analysis as Table 3 but for quintile portfolios instead of individual assets. The portfolios are value weighted within ranks based on their respective market betas estimated on the same method. We measure the performance in predicting the next-period OLS beta as in Equation (4).

Table 3, the $\hat{\beta}^{\text{SW}}$ method has the highest R^2 across indices, nevertheless, the values hardly surpass 10%, respectively 2% in the monthly period. The regression coefficients $\gamma_{\hat{\beta}}$ are smaller, yet always highly statistically significant. The RMSE and MAE of the shorter periods increase dramatically and in the monthly case, $\hat{\beta}^{\text{OLS}}$ even surpasses the threshold of one. These results suggest that asset pricing tests based on portfolio sorts will likely suffer from this instability. We are not aware of a comparable study for stocks, but since the yearly portfolios do not even attain the predictability of a universe of individual stocks, we expect that stock portfolios exhibit much higher predictability of their beta factors. At least with the ubiquitous market factor, the shorter horizon estimations are highly noisy and it is reasonable to assume it will be similar for other factors. Again, the CEV bias is an issue here, but it does not increase relative to the drop in R^2 compared to yearly betas in Table 3. It is a measure of an average bias pertaining to the pooled sample, which suggests that the poor performance on a shorter horizon is driven rather by individual outliers. This motivates more research not only on how to choose an appropriate market index but also into how to construct a universe of crypto assets that is relevant and more importantly resonates with expectations following decades of financial research.

Quarterly betas										
Index	Predictor	T	$1-\lambda$	γ_0	$\gamma_{\hat{\beta}}$	SD	$\gamma_{\hat{\beta}}/\lambda$	R^2	RMSE	MAE
BDMI	$\hat{\beta}^{\mathrm{OLS}}$	19870	0.318	0.557	0.302	0.000	0.443	0.094	0.392	0.266
	$\hat{\beta}^{\mathrm{SW}}$		0.146	0.412	0.481	0.000	0.563	0.139	0.334	0.239
	$\hat{\beta}^{\mathrm{VCK}}$		0.260	0.388	0.508	0.000	0.686	0.116	0.329	0.236
CCMIX	$\hat{\beta}^{\mathrm{OLS}}$	18214	0.337	0.644	0.179	0.000	0.270	0.033	0.449	0.326
	$\hat{\beta}^{\mathrm{SW}}$		0.151	0.576	0.262	0.000	0.309	0.040	0.395	0.298
	$\hat{\beta}^{\mathrm{VCK}}$		0.257	0.588	0.245	0.001	0.330	0.026	0.386	0.290
CRIX	$\hat{\beta}^{\mathrm{OLS}}$	19870	0.289	0.632	0.220	0.000	0.309	0.053	0.470	0.331
	$\hat{\beta}^{\mathrm{SW}}$		0.120	0.519	0.352	0.000	0.404	0.075	0.402	0.297
	$\hat{\beta}^{\mathrm{VCK}}$		0.223	0.508	0.370	0.000	0.476	0.062	0.391	0.286
SCI	$\hat{\beta}^{\mathrm{OLS}}$	19870	0.316	0.644	0.274	0.000	0.401	0.077	0.435	0.301
	$\hat{\beta}^{\mathrm{SW}}$		0.136	0.510	0.427	0.000	0.494	0.105	0.374	0.270
	$\hat{\beta}^{\mathrm{VCK}}$		0.250	0.503	0.443	0.001	0.591	0.083	0.369	0.266
Monthly betas										
Index	Predictor	T	$1-\lambda$	γ_0	$\gamma_{\hat{\beta}}$	SD	$\gamma_{\hat{\beta}}/\lambda$	R^2	RMSE	MAE
BDMI	$\hat{\beta}^{\mathrm{OLS}}$	26012	0.290	0.726	0.085	0.006	0.120	0.008	1.097	0.566
	$\hat{\beta}^{\mathrm{SW}}$		0.197	0.592	0.252	0.011	0.314	0.019	0.859	0.486
	$\hat{\beta}^{\mathrm{VCK}}$		0.297	0.581	0.266	0.013	0.378	0.017	0.845	0.469
CCMIX	$\hat{\beta}^{\mathrm{OLS}}$	24354	0.298	0.744	0.061	0.006	0.087	0.004	1.225	0.683
	$\hat{\beta}^{\mathrm{SW}}$		0.185	0.669	0.153	0.012	0.188	0.007	0.981	0.591
	$\hat{\beta}^{\mathrm{VCK}}$		0.284	0.651	0.176	0.013	0.246	0.007	0.955	0.562
CRIX	$\hat{\beta}^{\mathrm{OLS}}$	26012	0.281	0.786	0.078	0.006	0.108	0.006	1.255	0.714
	$\hat{\beta}^{\mathrm{SW}}$		0.162	0.710	0.166	0.011	0.198	0.009	1.013	0.621
	$\hat{\beta}^{\mathrm{VCK}}$		0.263	0.694	0.189	0.012	0.256	0.009	0.986	0.591
SCI	$\hat{\beta}^{\mathrm{OLS}}$	26012	0.296	0.813	0.078	0.006	0.111	0.006	1.185	0.661
	$\hat{\beta}^{\mathrm{SW}}$		0.181	0.695	0.210	0.011	0.256	0.014	0.946	0.566
	$\hat{\beta}^{\mathrm{VCK}}$		0.285	0.688	0.225	0.012	0.315	0.012	0.926	0.542

Table 5 **This table shows one-period ahead forecasting performance for quarterly and monthly betas.**

5. Conclusion

The emerging asset pricing literature in crypto faces several specific challenges. Besides relatively short history and quickly changing universe of assets, extreme volatility and illiquidity make the standard financial analysis difficult. Also, the current research in crypto has yet to agree on an appropriate representation of the market factor. We show that choosing a market risk factor is consequential for empirical results, at least in the context of estimating an asset's fundamental property, which is its market beta.

We test the predictability of crypto market betas and find that it is much lower, at least compared to recent results about the US stock market. Market betas are instrumental in asset pricing models as well as in hedging the market risk. Crypto investors face much higher time variation and errors-in-variables bias compared to the investors in the US stock market. The main takeaway is that while the US stocks exhibit predictability of about 55% for one-year-ahead market betas, historical betas in crypto explain at most 20% of the future market betas variation in S&P's BDMI index and less than 10% for other considered indices. This limits the out of sample hedging accuracy. Similarly to stock market results, we find that alternative methods that control for returns' outliers are better in predicting the OLS betas, than the naive OLS method itself.

Our findings point towards at least two direct avenues for further research. Firstly, we illustrate on a small sample that individual asset's market beta dynamics over time differ greatly. Potentially, there can be characteristics which would separate assets with low and high persistence of the market betas. Secondly, the asset pricing literature on crypto, which attempts to explain the excess returns, has so far discussed only cross-sectional patterns in characteristics rather than risk factor loadings of individual assets. This work outlines empirical difficulties that could be expected in crypto from the evidence of unstable market factor loadings in other risk factors as well. Thus inspection of methods improving beta estimation stability is likely necessary. Those issues are in fact exacerbated on quarterly and monthly horizons, which are typically used in such studies.

Acknowledgments

This work was supported by the Cooperatio Program at Charles University, research area Economics. Support from the Czech Science Foundation (project 20-17295S) is also highly appreciated.

References

Adrian T, Franzoni F (2009) Learning about beta: Time-varying factor loadings, expected returns, and the conditional capm. *Journal of Empirical Finance* 16(4):537–556.

Ang A, Hodrick RJ, Xing Y, Zhang X (2006) The cross-section of volatility and expected returns. *The journal of finance* 61(1):259–299.

Ang A, Liu J, Schwarz K (2020) Using stocks or portfolios in tests of factor models. *Journal of Financial and Quantitative Analysis* 55(3):709–750.

Bali TG, Engle RF, Murray S (2016) *Empirical asset pricing: The cross section of stock returns* (John Wiley & Sons).

Baur DG, Hong K, Lee AD (2018) Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money* 54:177–189.

Berentsen A, Schär F (2019) Stablecoins: The quest for a low-volatility cryptocurrency. *The economics of Fintech and digital currencies* 65–75.

Bianchi D, Babiak M (2021) A factor model for cryptocurrency returns. *Available at SSRN 3935934* .

Blume ME (1970) Portfolio theory: a step toward its practical application. *The Journal of Business* 43(2):152–173.

Blume ME (1971) On the assessment of risk. *The Journal of Finance* 26(1):1–10, URL `http://dx.doi.org/https://doi.org/10.1111/j.1540-6261.1971.tb00584.x`.

Bos T, Newbold P (1984) An empirical investigation of the possibility of stochastic systematic risk in the market model. *Journal of Business* 35–41.

Brenner M, Smidt S (1977) A simple model of non-stationarity of systematic risk. *The Journal of Finance* 32(4):1081–1092.

Bruner RF, Li W, Kritzman M, Myrgren S, Page S (2008) Market integration in developed and emerging markets: Evidence from the capm. *Emerging Markets Review* 9(2):89–103.

Carnero MA, Peña D, Ruiz E (2012) Estimating garch volatility in the presence of outliers. *Economics Letters* 114(1):86–90.

Chang BY, Christoffersen P, Jacobs K (2013) Market skewness risk and the cross section of stock returns. *Journal of Financial Economics* 107(1):46–68.

Chang I, Tiao GC, Chen C (1988) Estimation of time series parameters in the presence of outliers. *Technometrics* 30(2):193–204.

Conrad C, Custovic A, Ghysels E (2018) Long-and short-term cryptocurrency volatility components: A garch-midas analysis. *Journal of Risk and Financial Management* 11(2):23.

Corbet S, Lucey B, Urquhart A, Yarovaya L (2019) Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis* 62:182–199.

Dimson E (1979) Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics* 7(2):197–226.

Dutta A, Bouri E (2022) Outliers and time-varying jumps in the cryptocurrency markets. *Journal of Risk and Financial Management* 15(3):128.

Dyhrberg AH (2016) Bitcoin, gold and the dollar–a garch volatility analysis. *Finance Research Letters* 16:85–92.

Engle RF (2016) Dynamic conditional beta. *Journal of Financial Econometrics* 14(4):643–667.

Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. *Journal of financial economics* 33(1):3–56.

Fama EF, French KR (2004) The capital asset pricing model: Theory and evidence. *Journal of economic perspectives* 18(3):25–46.

Fama EF, MacBeth JD (1973) Risk, return, and equilibrium: Empirical tests. *Journal of political economy* 81(3):607–636.

Gebhardt WR, Hvidkjaer S, Swaminathan B (2005) The cross-section of expected corporate bond returns: Betas or characteristics? *Journal of financial economics* 75(1):85–114.

Grané A, Veiga H (2010) Wavelet-based detection of outliers in financial time series. *Computational Statistics & Data Analysis* 54(11):2580–2593.

Groenewold N, Fraser P (1999) Time-varying estimates of capm betas. *Mathematics and Computers in Simulation* 48(4-6):531–539.

Harvey CR, Liu Y, Zhu H (2016) … and the cross-section of expected returns. *The Review of Financial Studies* 29(1):5–68.

Häusler K, Xia H (2022) Indices on cryptocurrencies: an evaluation. *Digital Finance* 1–19.

Härdle WK, Harvey CR, Reule RCG (2020) Understanding Cryptocurrencies*. *Journal of Financial Econometrics* 18(2):181–208, ISSN 1479-8409, URL `http://dx.doi.org/10.1093/jjfinec/nbz033`.

Jegadeesh N, Noh J, Pukthuanthong K, Roll R, Wang J (2019) Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation. *Journal of Financial Economics* 133(2):273–298.

Jensen MC, Black F, Scholes MS (1972) The capital asset pricing model: Some empirical tests. *Capital Markets: Asset Pricing & Valuation* .

Kamps J, Kleinberg B (2018) To the moon: defining and detecting cryptocurrency pump-and-dumps. *Crime Science* 7(1):1–18.

Katsiampa P (2017) Volatility estimation for bitcoin: A comparison of garch models. *Economics Letters* 158:3–6.

Kelly BT, Pruitt S, Su Y (2019) Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics* 134(3):501–524.

Kristoufek L (2018) On bitcoin markets (in) efficiency and its evolution. *Physica A: Statistical Mechanics and its Applications* 503:257–262.

Kristoufek L (2021) Tethered, or untethered? on the interplay between stablecoins and major cryptoassets. *Finance Research Letters* 43:101991.

Kristoufek L, Vosvrda M (2019) Cryptocurrencies market efficiency ranking: Not so straightforward. *Physica A: Statistical Mechanics and its Applications* 531:120853.

Kukacka J, Kristoufek L (2020) Do 'complex'financial models really lead to complex dynamics? agent-based models and multifractality. *Journal of Economic Dynamics and Control* 113:103855.

Lewellen J, Nagel S (2006) The conditional capm does not explain asset-pricing anomalies. *Journal of financial economics* 82(2):289–314.

Li T, Shin D, Wang B (2021) Cryptocurrency pump-and-dump schemes. *Available at SSRN 3267041* .

Lintner J (1965) Security prices, risk, and maximal gains from diversification. *The journal of finance* 20(4):587–615.

Liu Y, Tsyvinski A (2021) Risks and returns of cryptocurrency. *The Review of Financial Studies* 34(6):2689–2727.

Liu Y, Tsyvinski A, Wu X (2022) Common risk factors in cryptocurrency. *The Journal of Finance* 77(2):1133–1177.

Mark M, Sila J, Weber TA (2020) Quantifying endogeneity of cryptocurrency markets. *The European Journal of Finance* 1–16.

Mishra DR, O'Brien TJ (2005) Risk and ex ante cost of equity estimates of emerging market firms. *Emerging Markets Review* 6(2):107–120.

Mossin J (1966) Equilibrium in a capital asset market. *Econometrica: Journal of the econometric society* 768–783.

Ramos S, Pianese F, Leach T, Oliveras E (2021) A great disturbance in the crypto: Understanding cryptocurrency returns under attacks. *Blockchain: Research and Applications* 100021.

Roll R (1977) A critique of the asset pricing theory's tests part i: On past and potential testability of the theory. *Journal of financial economics* 4(2):129–176.

Shah A, Chauhan Y, Chaudhury B (2021) Principal component analysis based construction and evaluation of cryptocurrency index. *Expert Systems with Applications* 163:113796.

Shahzad SJH, Bouri E, Roubaud D, Kristoufek L, Lucey B (2019) Is bitcoin a better safe-haven investment than gold and commodities? *International Review of Financial Analysis* 63:322–330.

Shanken J (1992) On the estimation of beta-pricing models. *The review of financial studies* 5(1):1–33.

Sharpe WF (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance* 19(3):425–442.

Shen D, Urquhart A, Wang P (2020) A three-factor pricing model for cryptocurrencies. *Finance Research Letters* 34:101248.

Trimborn S, Härdle WK (2018) Crix an index for cryptocurrencies. *Journal of Empirical Finance* 49:107–122.

Urquhart A (2016) The inefficiency of bitcoin. *Economics Letters* 148:80–82.

Urquhart A, Zhang H (2019) Is bitcoin a hedge or safe haven for currencies? an intraday analysis. *International Review of Financial Analysis* 63:49–57.

Vasicek OA (1973) A note on using cross-sectional information in bayesian estimation of security betas. *The Journal of Finance* 28(5):1233–1239.

Welch I (2019) Simpler better market betas. Technical report, National Bureau of Economic Research.

Welch I (2021) Simply better market betas. *Available at SSRN 3371240* .

Wooldridge JM (2015) *Introductory econometrics: A modern approach* (Cengage learning).

Zhang W, Li Y (2020) Is idiosyncratic volatility priced in cryptocurrency markets? *Research in International Business and Finance* 54:101252.

IES Working Paper Series

2022

1. Klara Kantova: *Parental Involvement and Education Outcomes of Their Children*
2. Gabriel Nasser, Doile de Doyle, Paulo Rotella Junior, Luiz Célio Souza Rocha, Priscila França Gonzaga Carneiro, Rogério Santana Peruchi, Karel Janda, Giancarlo Aquila: *Impact of Regulatory Changes on Economic Feasibility of Distributed Generation Solar Units*
3. Paulo Rotella Junior, Luiz Célio Souza Rocha, Rogério Santana Peruchi, Giancarlo Aquila, Karel Janda, Edson de Oliveira Pamplona: *Robust Portfolio Optimization: A Stochastic Evaluation of Worst-Case Scenarios*
4. Adam Kučera, Evžen Kočenda, Aleš Maršál: *Yield Curve Dynamics and Fiscal Policy Shocks*
5. Karel Janda, Ladislav Kristoufek, Barbora Schererova, David Zilberman: *Price Transmission and Policies in Biofuels-Related Global Networks*
6. Daniel Kolář: *Wealth Survey Calibration: Imposing Consistency with Income Tax Data*
7. Michal Hlaváček, Ilgar Ismayilov: *Meta-analysis: Fiscal Multiplier*
8. Salim Turdaliev, Karel Janda: *Increasing Block Tariff Electricity Pricing and the Propensity to Purchase Dirty Fuels: Empirical Evidence from a Natural Experiment*
9. Vojtěch Mišák: *Crime and weather. Evidence from the Czech Republic.*
10. Lukas Janasek: *Acquisition of Costly Information in Data-Driven Decision Making*
11. Josef Švéda, Jaromír Baxa, Adam Geršl: *Fiscal Consolidation under Market's Scrutiny: How Do Fiscal Announcements Affect Bond Yields*
12. Lenka Šlegerová: *How Is the Career Choice of a Medical Speciality Dependent on Gender Inequality in the Region*
13. Evgeniya Dubinina, Javier Garcia-Bernardo, Petr Janský*: The Excess Profits during COVID-19 and Their Tax Revenue Potential*
14. Ali Elminejad, Tomas Havranek, Zuzana Irsova: *People Are Less Risk-Averse than Economists Think*
15. Fan Yang, Tomas Havranek, Zuzana Irsova, Jiri Novak: *Hedge Fund Performance: A Quantitative Survey*
16. Eduard Baumöhl and Evžen Kočenda: *How Firms Survive in European Emerging Markets: A Survey*
17. Petr Jakubik, Saida Teleu: *Do EU-Wide Stress Tests Affect Insurers' Dividend Policies?*
18. Boris Fisera: *Exchange Rates and the Speed of Economic Recovery: The Role of Financial Development**
19. Jan Šíla, Michael Mark, Ladislav Krištoufek: *On Empirical Challenges in Forecasting Market Betas in Crypto Markets*

Univerzita Karlova v Praze, Fakulta sociálních věd
Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26
E-mail : ies@fsv.cuni.cz http://ies.fsv.cuni.cz