Accepted Manuscript

Asset pricing model uncertainty

Daniel Borup

PII: S0927-5398(19)30060-X
DOI: https://doi.org/10.1016/j.jempfin.2019.07.005
Reference: EMPFIN 1125

To appear in: Journal of Empirical Finance

Received date: 2 December 2017
Revised date: 11 June 2019
Accepted date: 10 July 2019


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Asset pricing model uncertainty

Daniel Borup

Aarhus University, Department of Economics and Business Economics, Fuglesangs Allé 4, Building 2622, 8210 Aarhus V, Denmark, and CREATES, Aarhus, Denmark. Email: dborup@econ.au.dk, Tel.: +45 87165134.

Abstract
This paper provides a unified calendar-time portfolio methodology for assessing whether returns following an event are abnormal which efficiently handles asset pricing model uncertainty and allows for time-varying alpha and factor exposures. The approach disciplines researchers' use of asset pricing factors and assigns a probability measure to the appropriateness of (dynamically) selecting a single model that best approximates the true factor structure or whether model averaging across an asset pricing universe is desired. It is applied to the long-horizon effect of dividend initiations and resumptions in the 1980 to 2015 period. Resulting post-announcement conditional abnormal returns are generally significant statistically and economically, which contrasts recent evidence, and exhibits a break in mean from positive until the mid-1990s and negative onwards. We document substantial time-variation in the dimensionality and composition of the factor structure in expected returns, which goes beyond what captured by conditional versions of the CAPM and Fama-French specifications. This also generalizes to a large panel of 202 characteristics-sorted portfolios.

Keywords: Abnormal returns, model uncertainty, time-varying risk, event study, calendar-time portfolio returns, dividend initiations

JEL Classifications: C18, G12, G14, G30, G35
Highlights

- A new calendar-time portfolio method for assessing abnormal returns is proposed.
- The method jointly handles asset pricing model uncertainty and time-varying coefficients.
- Abnormal returns following dividend initiations are generally negative and significant.
- They contrast those from CAPM and Fama-French models and recent literature.
- This is due to model misspecification and time-varying return factor structure.
Asset pricing model uncertainty

Abstract
This paper provides a unified calendar-time portfolio methodology for assessing whether returns following an event are abnormal which efficiently handles asset pricing model uncertainty and allows for time-varying alpha and factor exposures. The approach disciplines researchers’ use of asset pricing factors and assigns a probability measure to the appropriateness of (dynamically) selecting a single model that best approximates the true factor structure or whether model averaging across an asset pricing universe is desired. It is applied to the long-horizon effect of dividend initiations and resumptions in the 1980 to 2015 period. Resulting post-announcement conditional abnormal returns are generally significant, statistically and economically, which contrasts recent evidence, and exhibits a break in mean from positive until the mid-1990s and negative onwards. We document substantial time-variation in the dimensionality and composition of the factor structure in expected returns, which goes beyond what captured by conditional versions of the CAPM and Fama-French specifications. This also generalizes to a large panel of 202 characteristics-sorted portfolios.

Keywords: Abnormal returns, model uncertainty, conditional asset pricing, event study, calendar-time portfolio returns, dividend initiations

JEL Classifications: C18, G12, G14, G30, G35
1. Introduction

The purpose of an event study is to isolate the incremental impact of an event on price performance relative to established determinants of performance (Kothari and Warner, 2006). In long-horizon event studies on abnormal stock returns, a widely applied methodology is the calendar-time portfolio (CTP) approach. The method constructs at each time point, $t$, an event portfolio of the firms which have experienced the event in question within a fixed horizon prior to time $t$. Inference on the event effect is then based on a Jensen’s alpha procedure (Jensen, 1967; Jaffe, 1974; Mandelker, 1974) which amounts to evaluating the intercept in a time series CAPM or Fama-French regression, possibly including a momentum factor (Carhart, 1997).

Two most crucial problems with inference are, however, reliance on a single, possibly misspecified model of asset pricing and the imposition of constant coefficients (“alpha” and “betas”). First, the true asset pricing model is unknown, and perfect factor selection rarely happens in practice. In other words, model uncertainty is predominant. Moreover, investors may be concerned about different types of risk conditional on, e.g., the economic environment, such that the composition of the factor structure in returns is time-varying (Fama, 1998; Lyon et al., 1999; Johannes et al., 2014). Secondly, due to changes over time in the composition of the event portfolio or shifts in the surrounding economic environment, factor exposures are most likely time-varying (Fama and French, 1997; Boehme and Sorescu, 2002; Armstrong et al., 2013).

In this paper we introduce a unified framework which allows for time-varying alpha and betas as well as efficiently and dynamically handles asset pricing model uncertainty. We specify a dynamic linear model (West and Harrison, 1997) for the return-generating process where coefficients are defined as an endogenous latent process such that their values are inferred from returns directly. This avoids the underconditioning bias typically faced in conditional asset pricing models where the researcher is required to instrument coefficients by certain observables that are supposed to proxy investors’ information set (Cochrane, 2005; Lewellen and Nagel, 2006). The degree of time-variation in coefficients can either be fixed by the researcher or inferred simultaneously by data. Our framework also nests the specification of constant regression coefficients allowing for both conditional and unconditional analysis of the pricing of excess returns.
To handle asset pricing model uncertainty and explicitly address the “bad-model problem” of Fama (1998), we define an asset pricing universe of factor models as spanned by (linear combinations of) \( K \) candidate factors derived from the asset pricing literature. This universe nests conventional models typically employed in the literature, e.g. CAPM or the three-factor Fama-French model, and evaluates the pricing performance of each factor both individually and in conjunction with others. Accordingly, Fletcher (2018) empirically finds that better performing pricing models can be constructed from new combinations of the set of factors included in a set of conventional global asset pricing models, see also Bryzgalova (2016), Fama and French (2018), and Asness et al. (2018). Some factors may not be uniformly strong for pricing purposes at all times when assessed in this asset pricing universe, but it is crucial for obtaining valid inference on abnormal returns that relevant factors for pricing excess returns and reducing an omitted variable bias are accounted for in the asset pricing model. If these factors are ignored, and they are correlated with the included factors, it may lead to incorrect inference and even sign of the abnormality in a given return pattern.

In response, we specify a dynamic Bayesian framework on the basis of Dangl and Halling (2012) and Raftery et al. (2010) which infers from data (at each point in time) the support for a given factor and factor model. In the interest of efficiency and parsimony, those factors should be included only to the degree they matter. Most financial literature focuses, however, on two polar viewpoints. Based on empirical tests they tend to either reject an asset pricing model and discard it or vice versa. Acknowledging as Fama (1998) that such models are inevitably simplifications of reality while at the same time recognizing that complete confidence in a single asset pricing model may be rather extreme, even if data rejects a model, one ought to still account for the information in the model or factors to some degree (see also Pástor (2000), Pástor and Stambaugh (2000), and Avramov (2004)). In our framework, this degree is measured by posterior model probabilities which capture the time-\( t \) support for a factor or model. Hence, to efficiently utilize the entire asset pricing universe, we can base our model choice on the time series of those posterior model probabilities which allow for either dynamic selection of the “most correct” model or averaging over all model combinations. Employing posterior model probabilities as weights in model averaging, which penalize overspecified models with needlessly many factors through Occam’s razor effect, the procedure allows for an efficient way of accommodating the
'multidimensional challenge' of expected returns (Cochrane, 2011) for the use in event studies or asset pricing in general.

Our framework assesses pricing models simultaneously and, as a result, it is less prone to pre-testing or repeated analyses for several model specifications.\(^1\) In essence, our proposal is a more disciplined approach to asset pricing that avoids data-snooping, relative to the conventional practice of selecting some arbitrary factors for inclusion in the benchmark model, which leaves researchers much more freedom (Fama and French, 2018). Rather, the researcher needs to specify a broad asset pricing universe, from which our framework delivers pricing, robust to model misspecification. Our ultimate goal is, hence, not to come up with the true asset pricing model. Instead, we acknowledge the fact that we do not know the true asset pricing model but are able to identify its closest approximation as a function of a potentially large set of candidate asset pricing factors whose elements we do not know when to (potentially) apply. However, if the researcher has strong prior knowledge or desire about the inclusion of certain factors, it can easily be accommodated by fixing the particular factors’ inclusion probability to unity at all time periods beforehand. If the researcher wants a test of abnormality against, say, the CAPM, robust to model misspecification and omitted variable bias, one fixes the inclusion of the market risk premium and treats the remainder of the asset pricing universe of risk/mispricing factors as control whose relevance is determined and corrected for by our proposed framework. Besides its econometric benefits, our framework gives rise to interesting insights into the main drivers of the anomaly return pattern under investigation.

Recent contributions to the literature on event study methodology have primarily focused on the characteristics-based matching framework, see, e.g.,

\(^1\)Repeated analyses is a common choice of methodology in event studies when the authors are uncertain about model specification. Gregory (1997) and Boehme and Sorescu (2002), for example, implicitly recognize the presence of model uncertainty in long-horizon event studies by implementing several different model specifications and report the results arising from all the different specifications. However, while such a procedure highlights differences between methods, it also leaves inferences incomplete unless all models lead to the same conclusion. That is, it is not given how one should distinguish among or synthesize the conclusions derived from different model specifications. We suggest a disciplined solution, providing a single measure of abnormality.
Bessembinder and Zhang (2013) and Bessembinder et al. (2018), who substantially refine those methods. However, according to Kozak et al. (2018b) no good characteristics-sparse pricing model exists for the cross-section of expected stock returns, limiting the applicability of procedures that require ex-ante choosing a somewhat low-dimensional characteristic-based model. As argued above, our proposal adds a required level of sophistication to the Jensen’s alpha/calendar-time portfolio framework, improving its usefulness. By doing so, our method is applicable to all long-horizon event studies and (un)conditional asset pricing in general which we discuss further below. As noted by Kothari and Warner (2006), Bessembinder and Zhang (2013) and Bessembinder et al. (2018) it is valuable to continue to refine both methods and assess the consistency of results across methods, because neither the characteristics-based matching method nor the calendar-time portfolio method to event studies are perfect. We, therefore, apply the technique to the widely discussed hypothesis of long-horizon overperformance following dividend initiations and resumptions, which continues to draw attention in the literature (Hameed and Xie, 2018), and illustrate the type of economic insights our framework gives rise to. Consequently, we take a conditional view and examine the “chance proposal” of Boehme and Sorescu (2002) who attribute a significant event effect to pure chance (choice of sample period).

Applying our unified framework, we find a significant one-, three- and five-year conditional abnormal return in the 1980 to 2015 period for an equally-weighted portfolio. The abnormality of the one-year value-weighted event portfolio is confined to sub-periods whereas the evidence is significant over the full sample for the three- and five-year horizons. Our conditional view reveals a clear regime shift in mid-1990s, switching from weakly positive to negative abnormal returns, interestingly matching with increases in the relative importance of share repurchases as an alternative payout policy. It appears, thus, that investors have changed their idea about dividends from being positive to negative signals, possibly realizing that managers’ decision of dividend initiations is not necessarily a sign of good future performance, and that it might even be decoupled from firm fundamentals (Chetty and Saez, 2005). Seen over the full sample, event effects for both portfolio weighting schemes are mostly negative. This contrasts the general conclusion in recent literature (see, e.g., Bessembinder and Zhang (2013), Chen et al. (2014), and Bessembinder et al. (2018)) that event effects are largely positive, yet insignificant. Interestingly, our unified framework and conventional models
disagree notably from late-1990s and onwards and sometimes by more than 1% in the conditional monthly alpha. Their abnormal returns are even negatively correlated showing opposite signs and movements in periods related to economic crises. In conjunction with substantial improvements in empirical fit of our CTP-DMA methodology, this indicates that conclusions from conventional models may be biased due to model misspecification.

Moreover, the 1980-2015 period is associated with increasing dimensionality of expected returns, from about six-seven factors in the 1980-1990s to about eight-nine most recently. Underlying these numbers is a strong time-variation in the pertinence of specific factors. For instance, the event portfolio returns contain a large compensation for macroeconomic risk in crisis periods. This is consistent with recent findings in Boons (2016), who finds that macroeconomic risks are significant in excess of market risk and conventional Fama-French factors. Interestingly, even in the best case scenario where the (time-varying) “most correct” asset pricing model (in the sense of the highest posterior model probability) is known and available to the researcher in each period, one still puts faith in models that are at best 97% likely not the true ones. This underscores the appropriate use of the entire asset pricing universe via Bayesian model averaging and the inappropriateness of rejecting information contained in parts of the asset pricing universe which cannot be captured by a dynamic selection of a single, best model. To assess these implications in a general asset pricing context, we also apply our methodology to a large panel of 202 characteristics-sorted portfolios in the spirit of Giglio and Xiu (2018). We confirm the superior performance of our DMA procedure relative to DMS and conventional models. We also find supporting evidence for the fact that asset pricing models should contain quite a large number of factors (about eight on average), with time-variation in the pertinence of factors, as well as notable time-variation in their associated abnormal returns.

The remainder of the paper is laid out as follows. The following section positions our proposal in the most recent and related general asset pricing literature. Section 2 introduces the setting of the CTP long-horizon event study and presents our unified framework. We also show how to conduct inference on both alpha and asset pricing factors and models. Section 3 presents the event study application to dividend initiations and resumptions and Section 4 conducts the application to the large panel of characteristics-
sorted portfolios. Section 5 concludes.

1.1. Relation to general asset pricing literature

Treating model uncertainty and time variability of coefficients jointly in a unified framework, which nests several proposals introduced separately in the literature, also contributes to the very large and active research field on equity asset pricing. Instead of attempting a thorough review, we briefly relate our paper to some of the most recent contributions in the general asset pricing literature most closely related to our analysis.

A recent literature has focused on pitfalls in estimating and interpreting conventional factor models and that inference should be robust to model misspecification (Kan and Zhang (1999a,b), Shanken and Zhou (2007), Kan and Robotti (2008), Kan et al. (2010), Gospodinov et al. (2013), Burnside (2016), Bryzgalova (2016), and Gospodinov et al. (2017)). Existing literature considers mostly inference on pseudo-true coefficients under the presence of model misspecification. On the contrary, Feng et al. (2017) make inference on original coefficients by correcting for the model misspecification bias via a two-stage Lasso procedure that selects relevant asset pricing factors for excess returns and reducing an omitted variable bias. Our Bayesian alternative achieves similar ideas, but with a different objective. Their interest is in cross-sectional inference on the risk price of a new candidate factor. Our interest is in time series pricing and, primarily, an associated alpha. Moreover, we explicitly model time-varying coefficients. Despite the fundamental differences, our general logic is similar in that we acknowledge the presence of model uncertainty and that perfect model selection may be implausible requiring the researcher to specify a broad asset pricing universe from which we can get a reliable approximation of the true model.

Kelly et al. (2019) propose an intriguing Instrumented PCA (IPCA) approach. They treat factors as latent and instrument alpha and betas with firm characteristics. The factors are estimated with PCA as those components that line up with characteristics-driven exposures. If they do not exist, the characteristics-driven alpha will dominate. Our methodology differs from theirs in several ways. First, we retain interpretability of the driving risk/mispricing patterns of the excess return to be priced. We allow the researcher to determine the pertinence of a factor or factor model at a given point in time. This is relevant because a researcher evaluating a candi-
date anomalous return pattern would be interested in assessing not only its alpha, but also whether it reflects compensated risk and/or resembles an already documented anomaly. Such interpretability is generally lost in factors derived statistically (from, e.g., PCA). Secondly, the implementation of the IPCA requires the researcher to take a stance of investors’ information set by specifying the instruments of the coefficients which drive the time variability. As pointed out in Cochrane (2005), investors’ information is inherently unobserved for the econometrician, making conditional asset pricing models impossible to test. Indeed, Halling et al. (2018), find that betas are largely described by a yet unknown component. To avoid this issue, it is preferable to infer coefficients directly from information in returns themselves. Lewellen and Nagel (2006) develop an early contribution that acknowledges this point. They propose to estimate conditional CAPM betas from daily or weekly data within a month or quarter. The method by Lewellen and Nagel (2006) applies, however, only in the case where daily (or weekly) data is available on factors which may be restrictive given that many factors are constructed on the basis of low-frequent macroeconomic or accounting data. There is, too, no need for instrumenting betas (or alpha) in our framework. Thirdly, with no risk or mispricing interpretation of the identified IPCA factors, the interpretation of their IPCA betas is also reduced in that it is not clear which type of return premium they represent exposure to.

Barillas and Shanken (2018) propose a Bayesian framework for unconditional asset pricing. Like us, they also examine an asset pricing universe spanned by linear combinations of a set of candidate factors and allow for simultaneous comparison between all of the implied factor models on the basis of posterior model probabilities. We differ methodologically in that we, in addition to model uncertainty, jointly treat time-variation in alpha and betas, allowing for (but not requiring) conditional asset pricing. Moreover, they focus on Bayesian model selection for absolute asset pricing. Our framework allows the researcher to determine the appropriateness of selecting a single model or the need for Bayesian model averaging which utilizes relevant information in the remaining asset pricing universe. Consistent with an unconditional focus, model selection is conducted ex-post at the end of the sample in Barillas and Shanken (2018). We, in addition, show how to dynamically assess factor models, consistent with conditional asset pricing that maps directly to the decision-making of investors who conduct asset allocation based on conditional expectations. Lastly, as the authors also point out themselves, their
set of candidate asset pricing factors are somewhat limited. In fact, their preferred model exploits the maximum amount of factors possible suggesting that additional factors may be relevant. We find an average model size of seven-eight factors, derived from a greater asset pricing universe spanned by additional factors, which exceeds that of Barillas and Shanken (2018)’s six factors.

2. Methodology

We are in essence concerned about the average long-horizon abnormal return across \( N \) event firms after the occurrence of an event. According to the CTP procedure, we construct in each time period (months) a portfolio of firms that have experienced the event within \( h \) periods prior to the given month \( t \). In monthly long-horizon event studies, one usually fixes \( h \) as either 12, 36, or 60, representing a one-, three-, and five-year event horizon, respectively. The monthly event portfolio return, \( R_{p,t} \), is computed as the equally-weighted average of the event firms’ monthly returns or as the value-weighted average return to better replicate investors’ total wealth effects from holding the portfolio. Denote by \( R_{f,t} \) the risk-free rate and define \( R_{e,p,t} = R_{p,t} - R_{f,t} \), the portfolio excess return. The conventional CTP approach then dictates the relation

\[
R_{e,p,t} = \alpha + X_t \beta + \varepsilon_t, \tag{1}
\]

where \( X \) (of appropriate dimension) usually contains the market risk premium and the Fama-French factors (Fama and French, 1993), possibly along with the momentum factor of Carhart (1997), and \( \varepsilon_t \) constitutes an error term. Under the assumption that the factor model used in (1) is the true asset pricing model describing expected excess returns, the intercept, \( \alpha \), measures the average abnormal return to the event portfolio following the event. That is, it measures the average impact of the event. Under the null hypothesis of no event effects, the intercept should be zero. Numerous studies employ this method, for instance Boehme and Sorescu (2002), Moeller et al. (2004), and Bessembinder and Zhang (2013).

2.1. Calendar-time portfolio returns, time-varying coefficients, and model uncertainty

In this section, we present a CTP framework which jointly allows for time-varying coefficients and a large number of candidate asset pricing factors in
a parsimonious manner. Specifically, the framework utilizes the concept of dynamic model averaging (DMA) (see, e.g., Raftery et al. (2010)) and builds upon the work of Dangl and Halling (2012).

Let the filtration $\mathcal{F}_t$ be the $\sigma$-field which represents the time-$t$ conditioning information set which, in line with the CTP methodology, contains asset pricing factors and past returns. In general, it holds that

$$R_{p,t}^e = \mathbb{E}[R_{p,t}^e|\mathcal{F}_t] + \epsilon_t,$$

(2)

where $\mathbb{E}[R_{p,t}^e|\mathcal{F}_t]$ is the conditional expectation of the event portfolio excess return, henceforth referred to as the normal return of the event portfolio, and $\epsilon_t$ is an error term. The overarching model in (2) is the starting point for our analysis. The purpose of the event study is then to isolate the incremental impact of the event on stock returns in excess of this normal return. Given the existence of the risk-free rate, we model the normal excess portfolio return as a linear factor pricing model which is rooted in conventional asset pricing theory (Cochrane, 2005). Suppose the factor model contains at most $K \geq 1$ pricing factors which the researcher deem potentially relevant for measuring normal returns. Like other asset pricing analyses, we require those factors to be traded portfolio excess returns and, consequently, in case of non-traded factors (e.g., real consumption growth), we construct relevant factor mimicking portfolios. The theoretical foundation of such procedure is documented in Breeden (1979) and Huberman et al. (1987), and the asset pricing model is without an intercept term. Thus, by adding an intercept term to the relation in (2), we introduce a measure of an abnormal return following a given event. The framework we propose below is Bayesian in nature. As such, it represents a setting where investors recursively learn about both alpha, factor exposures and the relevance of asset pricing factors, gradually incorporating their beliefs to form the conditional expected returns in (2).

The return-generating process is specified as a dynamic linear model (DLM)\(^2\)

where $X_t$ is a $(1 \times K + 1)$ vector of factors (including a constant) and $\theta_t = (\alpha_t, \beta_t)'$ is a $(K + 1 \times 1)$ vector of coefficients (alpha and betas). A few remarks are worthwhile here. First, coefficients are specified as an endogenous latent process where time variability is allowed for by letting them follow a random walk. This tends to be superior in empirical applications (Dangl and Halling, 2012). Essentially, changes in alpha and betas are unpredictable, consistent with the suggestion in Boehme and Sorescu (2002). Moreover, the case of constant regression coefficients is nested in (4) if $Q_t$ is equal to zero for all $t$. This coefficient specification avoids the need of instrumentation. This is suitable because the investors’ information set is latent to the econometrician, making any instrumentation only an approximation of the true time variability in coefficients. Indeed, Halling et al. (2018) document that factor exposures are dominated by a yet unknown, time-invariant component not captured by known firm characteristics nor their industry relation. Secondly, the stochastic dynamics for the variation in coefficients in the style of (4) utilize information efficiently by keeping track of information in the past and weighting most recent exogenous shocks higher. It can be shown that the weights assigned to past information contained in returns decay exponentially with the decay being a function of the signal-to-noise ratio in returns (Bianchi, 2018). Hence, we differ from the proposal in Lewellen and Nagel (2006) who use only contemporaneous, higher-frequent information to infer betas. This also contrasts, for instance, a rolling least squares approach which assumes coefficients are constant in each estimation and weighted equally, as implemented in Bessembinder et al. (2018).\textsuperscript{4}

\textsuperscript{3}From an econometric point of view, the assumption of random walk coefficients may theoretically allow coefficients to drift arbitrarily high or low causing returns to be non-stationary. However, as pointed out by Dangl and Halling (2012), data at monthly frequency or higher tend to provide a remedy to the concerns of non-stationarity since the high frequency of observations drives dynamics of estimated coefficients rather than an increase in variance coming from the random walk assumption. Moreover, we show below how to model the time-variation of coefficients in a data-driven manner which avoids any unreasonable behavior. Evidently, in the empirical section all estimated coefficients behave “reasonably”.

\textsuperscript{4}In fact, a recursive least squares approach can be obtained as a nested case of this
the context of an event study, this is especially relevant because the characteristics of the event portfolio may change substantially due to the rolling nature of creating cohorts of event firms. This requires the factor exposures to be responsive and, importantly, to weight recent information highest since a firm usually enters the portfolio only once and for a limited, fairly short period of time (depending on the event horizon). Changing economic environments, such as periods of increasing market stress, may also require factor exposures to be flexible. We document the relevance of both instances in the empirical section below. Lastly, in many long-horizon studies, the event follows a period of unusual firm performance such that the event itself may impose gradual changes in firm characteristics following the event (Bessembinder and Zhang, 2013; Hameed and Xie, 2018), causing factor exposures to change as also empirically documented in Boehme and Sorescu (2002). Via (4), we obtain a time series of Jensen’s alpha, measuring the event impact at each time \( t \). We show in Section 2.2 formal Bayesian inference in this setting which enables an examination of abnormal returns over time in line with the proposal of Fama (1998): “one way to test whether [an] anomaly is real or the sample-specific result of chance is to examine a different sample period”. Notably, this also reduces reliance on ad hoc subperiod investigations otherwise resorted to in the literature.

The DLM is itself restricted to implementing the same set of factors for all time periods. That is, it is unable to directly capture model uncertainty. For this matter, we consider a set of candidate DLMs. We define each model by the inclusion and exclusion of a factor from the set of \( K \) candidate factors. We, thus, consider an asset pricing universe of \( 2^K \) candidate models for each time period, such that denoting a particular model and quantities associated with it by \( k = 1, 2, \ldots, 2^K \) the set of models are

\[
R_{it}^{(k)} = X_i^{(k)}\theta_i^{(k)} + \varepsilon_i^{(k)}, \quad \varepsilon_i^{(k)} \sim N(0, H_i^{(k)}),
\]

\[
\theta_i^{(k)} = \theta_i^{(k-1)} + \eta_i^{(k)}, \quad \eta_i^{(k)} \sim N(0, Q_i^{(k)}).
\]

Since this setup would be computationally infeasible and result in imprecise inference unless \( K \) is small, Raftery et al. (2010) construct an accurate specification, see Branch and Evans (2006), but this do not efficiently utilize information due to the assumption of constancy in the recursive window, nor is it sufficiently responsive to new information (Li and Yang, 2011).
approximation that enables estimation via a single run of a Kalman filter for each model in each time period. That is, there is no need for a Markov Chain Monte Carlo (MCMC) algorithm.\footnote{If the model space gets too big for computational feasibility, one can resort to a slightly augmented procedure as suggested by Onorante and Raftery (2016) and Risse and Ohl (2017). They propose a dynamic Occam’s window (Madigan and Raftery, 1994) which dynamically eliminates most of the low posterior probability models determined by some lower threshold level set by the researcher while still being able to resurrect them if needed.} The approximation involves two forgetting parameters, $\delta$ and $\lambda$, both in the interval $(0, 1]$. Here, $\delta$ is related to the (uncertainty in) estimation of the factor loadings in (5)-(6) while $\lambda$ is related to the model probability component in same expression. That is, $\delta$ controls coefficient time variation whereas $\lambda$ controls model uncertainty. The Kalman filter prediction of the factor exposures and their covariance matrix are given by

$$
\hat{\theta}_{t-1|t-1}^{(k)} = \hat{\theta}_{t-1|t-1}^{(k)'} + \Sigma_{t-1|t-1}^{(k)} X_t^{(k)} \left( H_t^{(k)} + X_t^{(k)} \Sigma_{t-1|t-1}^{(k)} X_t^{(k)'} \right)^{-1} \left( R_{p,t} - X_t^{(k)} \hat{\theta}_{t-1|t-1}^{(k)} \right),
$$

and

$$
\Sigma_{t|t-1}^{(k)} = \Sigma_{t-1|t-1}^{(k)} - \Sigma_{t-1|t-1}^{(k)} X_t^{(k)'} \left( H_t^{(k)} + X_t^{(k)} \Sigma_{t-1|t-1}^{(k)} X_t^{(k)'} \right)^{-1} X_t^{(k)} \Sigma_{t|t-1}^{(k)}.
$$

In periods with high stock market stress, factor exposures of firms could be expected to change more frequently with the converse being expected in periods with low market volatility. Dangl and Halling (2012) provide a very elegant solution by associating a grid of $\delta$s with each factor model specification. We can then obtain a time series of estimates of time variability...
in coefficients by performing Bayesian model averaging (BMA) over this grid at each time point. Consequently, let $\delta_j \in \{\delta_1, \ldots, \delta_d\}$ denote a certain choice of time variation of the betas, and let $M_k$ denote a certain choice (model) of the $K$ candidate factors. The total (i.e., choice of both model and $\delta$) posterior model probability is given by

$$ p(M_k, \delta_j | F_t) = p(M_k | \delta_j, F_t)p(\delta_j | F_t), $$

(11)

the Kalman filter model prediction equations are given by

$$ p(M_k | \delta_j, F_{t-1}) = \frac{p(M_k | \delta_j, F_{t-1})}{\sum_{l=1}^{2^K} p(M_l | \delta_j, F_{t-1})}, $$

(12)

$$ p(\delta_j | F_{t-1}) = \frac{p(\delta_j | F_{t-1})}{\sum_{l=1}^{d} p(\delta_l | F_{t-1})}, $$

(13)

and the associated updating equations are

$$ p(M_k | \delta_j, F_t) = \frac{p(M_k | \delta_j, F_{t-1})p(R_{p,t}^e | M_k, \delta_j, F_{t-1})}{\sum_{l=1}^{2^K} p(M_l | \delta_j, F_{t-1})p(R_{p,t}^e | M_l, \delta_j, F_{t-1})}, $$

(14)

$$ p(\delta_j | F_t) = \frac{p(\delta_j | F_{t-1})p(R_{p,t}^e | \delta_j, F_{t-1})}{\sum_{l=1}^{d} p(\delta_l | F_{t-1})p(R_{p,t}^e | \delta_l, F_{t-1})}. $$

(15)

In the particular case of $\lambda = \delta_t = 1$ for all $t$, the procedure reduces to a constant coefficient BMA procedure.

In terms of modeling $H_t^{(k)}$, we follow the suggestions of, among others, Raftery et al. (2010), Bengt and Halling (2012), and Byrne et al. (2018) and assume that $H_t^{(k)} = H^{(k)}$ for all $t$. Conveniently, if the prior on $H^{(k)}$ is inverse gamma, and if the prior on $\theta^{(k)}$ is normal, integrating the conditional density of excess portfolio returns over the range of $\theta^{(k)}$ and $H^{(k)}$ leads to a closed-form solution of the marginal likelihood, given as the density of a t-distribution. This also enables fast recursive computations. Alternatively, a time-varying error variance can be allowed.

The setup in (7)-(15) constitutes the Kalman filter procedure for estimation where the recursions are initiated with a prior on the model probabilities, $p(M_k, \delta_j | F_0)$, and on the coefficients for all models. In the implementation of the procedure, we use a burn-in period in order to train the model and hyperparameters as in Avramov and Chao (2006). We discuss this further in Section 3.3.
2.2. Bayesian inference

When conducting the event study, we are interested in the value of alpha above all to determine the sign and size of a potential event-induced abnormal return. Additionally, we may be interested in examining betas to assess the exposure to return-generating factors. In order to obtain these, it is useful to examine the implications of the posterior model probabilities obtained via (14) and (15) first.

2.2.1. Model and factor posterior probabilities

The metric

\[ p(M_k | \mathcal{F}_t) = \sum_{j=1}^{d} p(M_k | \delta_j, \mathcal{F}_t) p(\delta_j | \mathcal{F}_t) \]

(16)

measures the posterior model probability of model \( k \) after integrating out the effect of \( \delta_j \). It measures the relative support of a particular model, e.g., the three-factor model of Fama and French (1993), relative to the remainder of the asset pricing universe. In the spirit of Barillas and Shanken (2018), this is a simultaneous model-comparison framework that analyzes models jointly as opposed to classical pairwise test often seen in the literature (Vuong, 1989). This information is useful for several purposes. First, it informs about whether there exists a single best model that receives the majority of the support. In this case, inference on the event effect and the identification of main return drivers could be reasonably based on the outcome from the model with highest posterior model probability at each point in time. It should be noted that the composition of this model need not be the same over time, given the dynamic feature of our framework, which is documented in the empirical section below. We refer to this procedure as dynamic model selection (DMS). If the relative strength of the single-best model is weak, much relevant information is left in the remainder of the asset pricing universe. It is crucial to account for this information in the interest of obtaining valid inference which is robust to model misspecification and omitted variable bias. Inference should be based on a weighted average over all the models in the asset pricing universe where the weights are determined by

\[ \text{Note that DMS effectively treats the posterior model probability for the selected model as equal to unity, making an oracle-like assumption that there exists only one relevant and identifiable asset pricing model for each } t \text{ in the asset pricing universe specified by the researcher, ignoring potential model uncertainty within each } t. \]
the (trajectory of) posterior model probabilities. We refer to this procedure as dynamic model averaging.\(^7\) In the empirical section below, the strongest model receives at most 3.5% weight across the full sample, suggesting a need for DMA. This goes well with points made in Fama (1998):

“[p.291] The problem is that all models for expected returns are incomplete descriptions of the systematic patterns in average returns during any sample period. [p.285] All models show problems describing average returns. The bad-model problem is ubiquitous, but it is more serious in long-term returns.”

At first hand, these procedures may hint at potential issues with overfitting and lack of efficiency, and, therefore, a method that may favor the null hypothesis in the asset pricing study. However, we put forward two arguments that contradict this. First, it is important to note that for an equal prior on each model, the posterior model probabilities are largely determined by the likelihood function. Since they integrate over the entire parameter space, parsimonious models are preferred over complex models that include needlessly many factors, effectively employing Occam’s razor effect. That is, the likelihood will be highest for parsimonious models that use only the parts of the asset pricing universe required to provide a satisfactory fit of the excess returns. Since we define the asset pricing universe by all linear combinations of the \(K\) factors, many of the models considered are parsimonious allowing the procedure to favor those if deemed relevant by data. Secondly, only if certain factors are relevant for excess returns and for reducing an omitted variable bias, will they be assigned a non-zero weight. If a factor is highly correlated with another factor, and they share a signal for excess returns, our procedure will, by construction, share the total posterior probability assigned to the common signal between those two factors. That is, both factors get assigned a non-zero weight but in a fair proportion compared to the remainder of the asset pricing universe. Generally, these weights are rarely unity at all times, nor are they constant. Hence, as opposed to techniques that

\(^7\)Note that in principle we could minimize the omitted variable bias by including all variables at once in a kitchen-zink manner. However, with a large set of candidate factors and limited post-event observations, this is no viable solution due to less powerful inference. Importantly, a kitchen-zink procedure imposes a 100% inclusion probability for all factors at all points, which is unnecessary given our empirical findings below and distorts inference on alpha and eliminates information content in the time-varying inclusion probabilities inferred from data.
apply hard thresholding in that a factor is either included or not at all times, e.g., in Feng et al. (2017), we address the trade-off in efficiency by using only the information in factors when data tells us it matters to do so and in a manner that weighs its relevance. This is a means to efficiently handle the multidimensional challenge in Cochrane (2011).

Denote by \( \gamma_{t,j}^{(k)} \) the size (number of factors included) of the \( k \)'th model with degree of coefficient uncertainty \( \delta_j \) at time \( t \). Then we define the time-\( t \) expected (average) model size as

\[
\hat{\gamma}_t = \sum_{j=1}^{d} \left[ \sum_{k=1}^{2^K} \gamma_{t,j}^{(k)} p(M_k|\delta_j, \mathcal{F}_t) \right] p(\delta_j|\mathcal{F}_t).
\]

The average model size is an interesting output informing about the dimensionality of expected excess returns over time. It may also function as an informal sanity check of the specification of the asset pricing universe. If it takes a value close to the total number of candidate factors, it may indicate that the chosen asset pricing universe is underspecified and potentially cause of an unintentional omitted variable bias.

We may also be interested in inference on single factors and their relevance in the asset pricing universe. To that end, the \( l \)'th factor inclusion probability, \( \hat{\zeta}_l \), may be determined as

\[
\hat{\zeta}_l = \sum_{k=1}^{2^K} \mathbb{1}_{\{k \subseteq l\}} p(M_k|\mathcal{F}_t)
\]

with \( \{k \subseteq l\} \) indicating that the \( k \)'th model includes the \( l \)'th candidate factor. The trajectory of inclusion probabilities informs about the main drivers of excess returns over time. It allows us to examine whether the candidate anomaly return pattern under investigation resembles already documented anomaly patterns (such as momentum), or whether it is driven by risk compensation (coming from, e.g., real consumption growth). Depending on the specification of the asset pricing universe, this gives rise to valuable information about the components of expected excess returns.
2.2.2. Abnormal returns and factor exposures

To obtain estimates of the event effect (alpha) and the factor exposures (betas), model averaging dictates for \( t = 1, \ldots, T \) recursive identification as

\[
\hat{\alpha}_t = \sum_{j=1}^{d} \left[ \sum_{k=1}^{2K} \alpha_{t|j}^{(k)} p(M_k | \delta_j, F_t) \right] p(\delta_j | F_t),
\]

\[
\hat{\beta}_t = \sum_{j=1}^{d} \left[ \sum_{k=1}^{2K} \beta_{t|j}^{(k)} p(M_k | \delta_j, F_t) \right] p(\delta_j | F_t),
\]

performing model averaging over both the uncertainty surrounding choice of pricing factors and degree of time variation in coefficients.\(^8\)

Formal inference on the event effect at a given point in time dictates the null hypothesis \( H_0: \alpha_t = 0 \) against the alternative hypothesis \( H_1: \alpha_t \neq 0 \) for some fixed \( t \). Hypothesis testing can be conducted on the basis of standard Bayes factors comparing a restricted model with zero intercept at time \( t \) to a model with unrestricted intercept at the same point in time without requiring this restriction to be imposed at any other time. Thus, a testing methodology, which determines the time-\( t \) probability that the intercept is zero, given the information set available up to that point in time, is based on the recursive Bayes factor

\[
BF_t = \frac{p(R_{p,t}^1, \ldots, R_{p,t}^e | F_t, H_0)}{p(R_{p,t}^1, \ldots, R_{p,t}^e | F_t, H_1)},
\]

such that, (see e.g., Robert (2007) and Bianchi (2018))

\[
p(H_0 | F_t) = \left( 1 + \frac{p(H_1 | F_0)}{p(H_0 | F_0)} BF_t \right)^{-1},
\]

given that the prior structure is common for the null and alternative model. With equal prior over the null and alternative hypothesis, \( p(H_0 | F_0) = p(H_1 | F_0) \),

\(^8\)The posterior distribution of coefficients admits a multivariate t-distribution, \( \mathcal{T}_{n_t - 1} (\hat{\theta}_t, \Sigma_t | F_t) \), where \( n_t = n_{t-1} + 1 \) is the degrees of freedom. Inference on coefficients is, thus, more robust to extreme realizations of exogenous shocks due to fatter tails compared to a normal distribution in the beginning of the sample when less information is available.
the posterior probability of the null hypothesis $p(H_0|F_t)$ is obtained as $BF_t/(1+BF_t)$.

This probability can broadly be interpreted as a frequentist $p$-value of the null hypothesis, but it differs fundamentally since it automatically includes Occam’s razor effect in that models with needlessly many factors that do not deliver improved pricing ability are penalized. The value $BF_T$ provides the means of inference over the full sample capturing the cumulative evidence against or in favour of the null hypothesis over all time periods.

3. The long-horizon effect of dividend initiations and resumptions

We now apply the unified CTP framework to the lively debated post-event stock performance following dividend initiations and resumptions. In a seminal article, Michaely et al. (1995) find a significant positive abnormal return of 24.8% for three years following dividend initiations in the period of 1964-1988 applying the characteristics-based matching (BHAR)

10

method using as benchmark the return on an equally-weighted market index. Boehme and Sorescu (2002) find supporting evidence when applying the BHAR method using dimensional matching on size and book-to-market ratios to construct benchmark returns. They document, in contrast, a disappearing abnormal return when applying the conventional CTP approach leading to the argument that the abnormal performance is attributed to pure chance (choice of sample period). A recent study by Chen et al. (2014) supports the “chance proposal” while controlling for the possibility of a negative bias arising from the presence of initial public offerings (IPOs) in the event sample. As a response to improving additional statistical rigour in event study methodologies, Bessembinder and Zhang (2013) propose a modified union of the BHAR and CTP approach and find that it leads to positive, yet insignificant abnormal returns following dividend initiations. Kolari et al. (2018) methodologically criticize Bessembinder and Zhang (2013) and find contrasting evidence in terms of

9 $BF_t$ may be conveniently computed using the Savage-Dickey density ratio and the fact that the prior structure is common for the null and alternative model. The ratio requires only computation under the alternative model, see, e.g., Bianchi (2018), and benefits from a closed-form expression of the marginal likelihood.

10 The buy-and-hold abnormal return for the $i$th firm with an event horizon of $T$ months is defined as the difference between the buy-and-hold realized return on the $i$th firm and the buy-and-hold return of a benchmark firm or portfolio supposed to capture the expected return over the event horizon. The event effect is then measured as the sample average of all event firms’ buy-and-hold abnormal returns.
positive and statistically significant long-horizon effects (see Bessembinder and Zhang (2017) for a response) confirming that long-run price effects are sensitive to choice of methodology. Bessembinder et al. (2018) document supporting evidence for the insignificant results in Bessembinder and Zhang (2013) in a novel characteristics-based matching event methodology. This overall ambiguity in findings motivates an application of our method proposed above to the long-horizon performance following dividend initiations and resumptions.

3.1. Events sample

We obtain the events sample following the lines of Michaely et al. (1995), Boehme and Sorescu (2002), and Bessembinder and Zhang (2013). For the 1980 to 2010 time period, we identify dividend initiations as the first cash dividend in the lifetime of a firm from the Center of Research in Security Prices (CRSP) database. We stop the events sample at year 2010 in order to ensure that all post-event windows may be up to five-years length. In addition, we identify dividend resumptions as the first cash dividend paid by a firm subsequent to a dividend omission with a hiatus of at least five years. We require that the firms have been listed in CRSP for a minimum of two years and have not experienced an IPO in two years before the dividend announcement, following Kale et al. (2012). The firms must be U.S. based, traded on the NYSE, Nasdaq, or AMEX exchanges, and be a common stock (share code 10 and 11). Lastly, the cash dividends must have a monthly, quarterly, semi-annual, or annual frequency. We also include the ones with an unspecified payment frequency since most of these have a quarterly payment frequency. In total, we identify 1016 event firms composed of 934 dividend initiations and 82 dividend resumptions. The distribution of events over the sample period is depicted in Figure 1.

We note that the yearly frequency of events decreases from mid-1990s to 2002, consistent with the disappearing dividend phenomenon (Fama and French, 2001), before surging in 2003-2005 which aligns well with the introduction of the Jobs and Growth Tax Relief Reconciliation Act of 2003. This act effectively reduced the tax rate on corporate dividends to 15%. It also aligns well with the timing of a correction to the SEC Rule 10b-18. The rule was released in 1982 to clarify circumstances under which corporations could make share
repurchases without being subject to dividend taxation, causing firms to adjust their payout policies, which caused changes to the informational content in dividend payments (Boudoukh et al., 2007). The rule was then altered in 2003 by requiring firms to disclose more detailed information about certain types of share repurchases. Chetty and Saez (2005) find that the increase in dividend initiations in this period is solely due to the tax cut and are not confounded by other factors that might have influenced the payout decision, including a “relabelling” of share repurchases as dividends after the reform and firm fundamentals.

3.2. Set of candidate asset (mis)pricing factors

Navigating in the zoo of asset pricing factors (Cochrane, 2011) is not a trivial task. Indeed, Harvey et al. (2016) identify as much as 316 factors published in leading journals and supposed to explain the cross-section of average returns. For many of those factors it is not clear whether they are driven by risk or common sources of mispricing as they are mostly observationally equivalent and both can belong in the pricing model, as discussed in Kozak et al. (2018a). Behavioral ideas (e.g., liquidity demands, distorted beliefs, etc.) result in mispricing only if they represent factor risk because otherwise they would not exist due to arbitrageurs exploiting them. Fortunately, this complication is not fatal in event studies. The purpose of an event study, no matter the method, is to isolate the incremental impact, relative to the controls, of an event on price performance. As also pointed out in Kothari and Warner (2006) and Loughran and Ritter (2000), it is a relevant empirical question to ask whether a potential anomaly pattern is distinct from other anomaly patterns already established. Since the price performance associated with size, book-to-market, and other characteristics is applicable to all stocks, not just the sample of event firms, the performance following the event itself must be different from those to be considered anomalous. From a methodological point of view, an event study methodology should, therefore, control for the relevant return patterns common to all stocks, regardless of a risk or mispricing interpretation.

Similarly in spirit to Giglio and Xiu (2018), we construct on this basis a set of candidate asset pricing factors combining macroeconomic, non-equity, and equity related factors which represent both common sources of mispricing or sources generally accepted to be true risk. Specifically, we include the well-known real consumption growth factor (CONS) (Breeden,
from the consumption-based asset pricing framework. Secondly, we include macroeconomic and non-equity factors proposed primarily in the seminal article by Chen et al. (1986), namely industrial production growth (IND), unanticipated inflation (INF), term spread (TMS), default risk premium (DFR), and default yield spread (DFY). Thirdly, we add equity-related factors, specifically, the prominent factors of the Fama-French five-factor model (Fama and French, 1993, 2015); market risk premium (MKT), small-minus-big (SMB), high-minus-low (HML), robust-minus-weak (RMW), and conservative-minus-aggressive (CMA). Additionally, we add momentum\textsuperscript{11} (MOM), betting-against-beta\textsuperscript{12} (BAB), illiquidity\textsuperscript{13} (ILLIQ), and quality-minus-junk\textsuperscript{14} (QMJ). Hence, the set of candidate factors includes the commonly used three factors of Fama and French (1993) and the momentum factor of Carhart (1997), thus facilitating comparison with other studies implementing conventional CTP models.

This amounts to a total of 15 different candidate pricing factors somewhat overlapping and extending the set in Harvey and Liu (2018) and Barillas and Shanken (2018), comprising a model space of $2^K = 32,768$ potential models each associated with $d$ degrees of time variation in coefficients. For all factors, we obtain a time series of monthly observations for the 1980 to 2015 period. The non-traded factors, INF, IND, and CONS, are transformed to traded factors using mimicking portfolios following the procedures in Jagannathan and Wang (2007) and Malloy et al. (2009). We obtain a monthly time series of per capita real consumption growth using the National Income and Product Accounts (NIPA) tables from 1959 to 2015.\textsuperscript{15} Data on the monthly production growth series and inflation is obtained from the Federal Reserve Bank of St. Louis database. We construct unanticipated inflation similarly to Chen et al. (1988) and Liu and Zhang (2008). Subsequently, we regress (including an intercept and macroeconomic controls, see Eckbo et al. (2000)) real consumption growth, industrial production growth, and unanticipated

\textsuperscript{11}See Carhart (1997).
\textsuperscript{12}See Frazzini and Pedersen (2014).
\textsuperscript{13}See Pástor and Stambaugh (2003).
\textsuperscript{14}See Asness, Frazzini and Pedersen (2019).
\textsuperscript{15}NIPA Table 2.8.5 provides the nominal personal consumption on non-durables and services. NIPA Table 2.8.4 provides the price deflators and NIPA Table 2.6 provides population size.
inflation on the excess returns of 30 industry portfolios obtained from Kenneth French’s data library, which provides the mimicking portfolio weights (in a least square sense). Consequently, the mimicking factor portfolio returns are these normalized weights multiplied by the excess returns of the benchmark portfolios. Data on TMS, DFR, and DFY is obtained from an updated Welch and Goyal (2008) data set. Data on the remaining (traded) factors is obtained from the respective authors’ websites. For elaboration on the factor constructions, we refer the reader to the specific references. Summary statistics on the final candidate asset pricing factors are reported in Table 1.

3.3. Robustness of parameter and prior choices

In order to implement the procedure outlined in Section 2, we need to elicit priors for the models, coefficients, and error variance along with a choice of range and granularity of the grid of $\delta$. For the latter, we choose $\delta \in \{0.97, 0.98, 0.99\}$ in the conditional analysis such that $d = 3$. We evaluated the results with a choice of granularity of $d = 7$, but results do not change notably. We use a non-informative model prior where all model combinations and degrees of time variation in the coefficients are equally likely, $p(M_k|F_0) = 1/2^K$ and $p(\delta_j|F_0) = 1/d$ such that $p(M_k, \delta_j|F_0) = 1/(d2^K)$ for all $k$ and $j$ consistent with Dangl and Halling (2012). Moreover, we use a proper prior on the coefficients following the suggestions in Barillas and Shanken (2018) with $\theta_{0,j}^k = N(0, g_{n_k})$ for all $k$ and $j$ where $n_k \in \{1, 2, \ldots, K\}$ is the number of included variables in the $k$’th model. We use a burn-in training sample of three-years’ observations to train $g$ which is common in Bayesian asset pricing analyses (see, e.g., Avramov and Chao (2006)).\footnote{Our choice of grid for values of $g$ is chosen to ensure a priori plausible deviations of alpha from zero. This is to avoid concerns of Bartlett’s paradox where posterior model probabilities favor a restricted, null model despite substantial alpha deviations from zero due to them being even further from the values envisioned under the unrestricted, alternative model. Empirically, we also found that very large values for $g$ perform worse based on a comparison of marginal likelihoods.} In this way, we put very little subjective information into the start of the procedure and initiate it under the null hypothesis of zero alpha. The prior on $H^k$ is inverse-gamma for all $k$, similar to Dangl and Halling (2012). We also implemented

\[\text{Insert Table 1 about here}\]
an uninformative natural conjugate $g$-prior specification of the Zellner (1986) type with $g = T$ as Nonejad (2017) and Byrne et al. (2018), and we found that the main conclusions are largely robust but somewhat affected primarily in the first part of the sample and mainly for factor exposures. We attribute this to the prior for the coefficient covariance matrix which is given by the $g$-scaled OLS estimate of the variance in coefficients over the entire sample. With only a limited length of the event portfolio return series, this $g$-prior seems to dominate too much initially before the Bayesian procedure is able to pick up empirical patterns in data. Finally, we set $\lambda = 0.98$ which is consistent with a case where model change is allowed, but not in a too rapid manner. The conclusions, though, do not materially change for $\lambda = 0.95$ and $\lambda = 0.99$.

3.4. Empirical results

For comparative purposes, we report in addition to the results from our unified framework also results applying conditional versions of conventional CTP models:

1. **CTP-DMA**: Calendar-time portfolio approach utilizing dynamic model averaging and all candidate factors.
2. **CTP-DMS**: Calendar-time portfolio approach utilizing dynamic model selection and all candidate factors.
3. **CTP-CAPM**: Conditional CAPM with the MKT factor.
4. **CTP-FF3**: Conditional three-factor Fama-French model with the MKT, SMB, and HML factors.
5. **CTP-FF4**: Conditional four-factor Fama-French-Carhart model with the MKT, SMB, HML, and MOM factors.
6. **CTP-FF5**: Conditional five-factor Fama-French model with the MKT, SMB, HML, RMW, and CMA factors.

In our analysis below, we fix the inclusion probability of the MKT to unity for all periods.$^{17}$ This is consistent with Barillas and Shanken (2018) and is

$^{17}$This reduces the model space to $2^{K-1} = 16,384$ combinations for each $d$. 

24
motivated by its unique role in portfolio analysis and equilibrium pricing of assets (Sharpe (1964), Lintner (1965), Mossin (1966), Merton (1973)). It also facilitates straight comparison of the CTP-DMA and CTP-DMS procedures to the CTP-CAPM, in that they differ precisely in the use of controls.

We report results for both equally- and value-weighted portfolios (EW and VW, respectively) where the monthly value-weights are constructed from prior month’s market capitalization obtained from the CRSP/Compustat merged database. The weights are rebalanced every month to reflect the changing composition of the portfolio. Considering both types of portfolios is relevant for economic as well as methodological reasons. First, value-weighting better reflects investors’ wealth effects from holding the portfolio. Secondly, event studies and general asset pricing studies tend to find different results, depending on the choice of portfolio (Boehme and Sorescu, 2002; Plyakha et al., 2016). Fama (1998) and Fama and French (2015) posits, indeed, that value-weighting mitigates problems with model misspecification, since it assigns a higher (lower) weight to larger (smaller) firms for which the problem is likely to be less (more) severe. Given our focus on model specification, it is thus interesting to examine the impact of model misspecification on the results from both EW and VW portfolios.

To examine the relative evidence for each model specification, Table 2 reports the empirical fit of the CTP-DMA relative to CTP-DMS and conventional models as measured by logarithmic Bayes factors. An advantage of using Bayes factors, as noted above, is their penalization of highly parametrized models that do not deliver improved empirical fit. In all but four cases, the evidence in favour of the CTP-DMA relative to all other models is either strong or very strong, in the wording of Kass and Raftery (1995). By corollary, the CTP-DMA provides a much more accurate characterization of the factor structure in the event portfolio returns.

---

18 We also completed the application without this requirement, and results are almost quantitatively identical due to MKT being chosen with probability one in close to all periods by DMA/DMS procedures.
19 According to the rule-of-thumb interpretation rules of Kass and Raftery (1995), values below 1 are not worth more than a bare mention, values between 1-3 are positive in favour of the CTP-DMA, values between 3-5 are strongly in favour of the CTP-DMA, and values greater than 5 are very strongly in favour of the CTP-DMA.
3.4.1. Empirical abnormal returns

Table 3 provides results on average conditional abnormal returns, inference, and average model sizes for one-, three-, and five-year post-event horizons for the dividend initiation and resumption event sample over the period of 1983-2015.  

For the EW portfolios, the conditional versions of the conventional models (CTP-CAPM, CTP-FF3, CTP-FF4, and CTP-FF5) show significant abnormal returns over the full sample, which are positive on average and economically large, across all horizons. The event effect is, though, decreasing as the horizon increases. These findings are consistent with Boehme and Sorescu (2002). The conditional alphas from the CTP-DMS, which dynamically selects the most correct model in each time period based on the posterior model probabilities, and CTP-DMA, which averages over the alphas across all factor models based on their relative posterior model probabilities, are generally much smaller. They disappear on average at five year horizons for both CTP-DMS and CTP-DMA and at three year horizons for CTP-DMA only. This disappearance is, however, a consequence of a major intertemporal difference in abnormal returns which can be seen from Figure 2 that depicts alpha with a three-year event horizon for CTP-DMA and conventional models.

Interestingly, the upper figure shows that abnormal returns are positive and stable from 1983 to 1988, zero and stable from 1989 to 1992, after which they are generally negative and less stable. At most time points, the event effect is

20We have run standard unconditional FF3 and FF4 CTP regressions as in Bessembinder and Zhang (2013) and confirmed their findings of small, insignificant average abnormal returns on EW portfolios. The results were similar for the VW portfolios. We have also run WLS regressions with the square root of number of firms contained in each month as weights. Such a procedure may be appropriate if the study concerns selective management events (such as dividend initiations, initial public offerings, or mergers and acquisitions) where the selection is partly driven by industry- and economy-wide conditions or stock misvaluations, because the events may be bunched in calendar time. The resulting abnormal return would then possibly be concentrated in certain periods. In the present paper, year 2003 and 2004 (cf. Figure 1) contain more events relative to remaining months. The WLS procedure should reveal any bunching effects, but the results were remarkably similar to the OLS results, providing no consistent support for bunching effects.
non-zero, but the positive effects in the first part of the sample cancel out on average the negative effects in remainder of the sample. On the basis of this analysis, it appears as if investors have changed their idea about dividend signalling from being positive to negative. It goes well in hand with the idea that the information content in dividend payments is changing over time, as documented in Goyal and Welch (2003) and Boudoukh et al. (2007). One explanation might be that investors have realized that managers’ decision of dividend initiations is not necessarily a sign of good future performance (it might, indeed, be the opposite, as profitability tends to decrease in the years subsequent to the announcement (Grullon et al., 2002)), and that it might even be decoupled from firm fundamentals. This is supported by Chetty and Saez (2005) who documents that the dividend initiation decision in the 2003-2004 period was solely driven by above-mentioned tax reform on dividend income and amplified by self-interest of managers. Additionally, they posit that shareholders’ awareness of severe accounting fraud in the 2001-2003 period might also have pressured management to pay out dividends.

Another explanation may arise from the increasing importance of share repurchases. Interestingly, the timing of the shift of sign in the abnormal return aligns perfectly with the empirical findings in Farre-Mensa et al. (2014) which show that the preferred payout policy shifted from being dividends before the mid-1990s to being share repurchases after the mid-1990s and until today. Share repurchases are now considered by executives the most efficient way of paying out money to investors, partly due to a preferable taxation scheme for investors relative to dividends and partly because of its flexibility (Brav et al., 2005). Moreover, investors are aware of the possibility for firms to pay out using share repurchases as an alternative to dividends (Grullon et al., 2002). Brav et al. (2005) find, additionally, that 80% of surveyed executives

\[\text{\textsuperscript{21}}\] In an earlier version of this paper, we conducted a (frequentist) test of change in the mean alpha based on a wild bootstrap procedure. The test rejected for all portfolios and event horizons, estimating the break for the equal-weighted three-year event horizon at 1994:3 with a positive (negative) mean in the first (second) part of the sample.

\[\text{\textsuperscript{22}}\] There seem to be no complete agreement why firms did not prefer share repurchases before the mid-1990s. One argument is that corporations were simply wrong (Grinblatt and Titman, 1998) and spent considerable time in learning that paying out cash through repurchases would not result in IRS taxing repurchases at the ordinary income tax rate (like dividends). Another is that there were too high a risk for violating antimanipulative provisions of the Securities Exchange Act of 1934 (Grullon et al., 2002).

27
make repurchase decisions after all their other investment plans are determined. Only a third of the executives responded that they make dividend decisions after investment plans are determined. The difference is statistically significant. This indicates that executives might pay out using share repurchases when they consider their own firm the most profitable investment available, whereas dividend payouts may occur before pursuing all profitable investment opportunities. In light of these payout policy characteristics, our findings indicates that when dividend payments were firms’ preferred payout policy, investors’ perceive dividend initiations as a positive signal. However, when firms and investors realize the appealing features of share repurchases relative to dividend payments, from the 1990s and onwards, they perceive dividend initiations as a negative signal. This can occur, since investors evaluate the choice of dividend initiations relative to the, possibly preferred, choice of share repurchases, and hence decrease their valuation of the firm, see also Hartzmark and Solomon (2019). No matter the explanation, it is clear that in order to extract this change in investor perception or information content, it is a necessity to allow for a time-varying alpha that quickly learns, in real-time, changes to the economic environment jointly with time-varying factor exposures and relevance which we facilitate in CTP-DMA and CTP-DMS.

The lower part of the figure reveals stark contrasts between the CTP-DMA abnormal returns and those obtained from conventional models pointing towards a severe issue of omitted variable bias, especially in the second part of the sample. From late 1990s and onwards, the trajectory of alpha is often-times negatively correlated between CTP-DMA alpha and all conventional models. At times, the difference is well above 1%. For instance, in the wake of the 2001 NBER recession conventional models pointed to a positive event effect, whereas CTP-DMA suggests a considerable negative effect. In the most recent recession, conventional models’ abnormal return decreased substantially to negative terrain, whereas CTP-DMA’s abnormal return increased from negative terrain to zero. The implication is that the significance of the different models’ abnormal returns is derived from largely different reasons. Conventional methods generally deem dividend initiations to be a positive event at most times, but the effect might have been negative during the most recent recession. On the contrary, CTP-DMA assesses the event effect to being positive only in the 1980s from which on the effect is generally negative. Overall, it stands out that choosing among conventional models
as the CAPM, FF3, FF4, or FF5 has little influence on the conclusions and measurement of event effects. However, all those models appear to be facing a severe omitted variable bias, causing misleading conclusions.\textsuperscript{23}

For the VW portfolio, all methods tend to agree on the sign of the average effect but not the magnitude nor significance.\textsuperscript{24} For a one-year horizon, the conditional event effect is significant for conventional models, but its average is economically insignificant. This contrasts the findings for CTP-DMA, which shows an economically significant negative event effect, but seen over the full sample it is insignificant. It is important to note that this does not imply that it is insignificant at all times. Figure 3 shows the cumulative evidence for no event effect based on (21) and (22).

During the late-1990s and up to the beginning of 2001 NBER recession, abnormal returns appeared significant by becoming significantly negative. Figure 4 depicts the CTP-DMA abnormal return for the VW portfolio and a three-year event horizon as well as the abnormal returns obtained from conventional models. The conclusions are analogue to those of the EW portfolio in that abnormal returns from conventional models are very different from (and generally negatively correlated with) CTP-DMA abnormal returns since mid-1990s.

Collectively, these findings strongly support the conditional nature of our analysis with a time varying alpha in (5) and, importantly, underscore the relevance of asset pricing model uncertainty for measuring abnormal returns. The discrepancy between CTP-DMA and CTP-DMS is less pronounced, but the difference in average conditional alpha remains noteworthy, nonetheless. This may arise from lack of empirical support in a single model, even the best

\textsuperscript{23} Relatedly, Bessembinder and Zhang (2013) documents that firm characteristics like idiosyncratic volatility, illiquidity, momentum and capital investment differ between event firms and firms matched on size and the book-to-market ratio, and that these differences vary over the event window. Our findings are consistent with Bessembinder and Zhang (2013) in that there are other variables determining expected returns to event firms than size and the book-to-market ratio, driving a wedge between conventional models and our proposals.

\textsuperscript{24} That average event effects disappear in conventional models when applying a VW portfolio as opposed to an EW portfolio is also found in Boehme and Sorescu (2002).
one. We analyze the factor composition of the CTP-DMS and CTP-DMA approach below, examining what causes these differences across models.

3.4.2. Empirical asset pricing model uncertainty

To get an overview of the relevance of each factor in our asset pricing universe, Table 4 reports average inclusion probabilities, $\hat{\zeta}_l$, $l = 1, \ldots, K$, from (18), of the candidate factors across the sample period.

For the EW portfolio, SMB is close to globally relevant in that it is included in almost all models at all times. This contrasts the case of the VW portfolio which reduces the SMB inclusion probability substantially. This seems natural given the construction of the portfolios. Across the macro space, real consumption growth seems to be the most relevant factor, especially at longer event horizons. Industrial production growth achieves, on the contrary, relatively little importance on average. There is some support for the conventional Fama-French models and the MOM in that they have a relatively high inclusion probability, mostly for the EW portfolio. These are, however, far from unity and are attached with a noticeable standard deviation, indicating they do not matter at all times, which is otherwise imposed by conventional models. Generally, most factors’ inclusion probability is non-negligible, indicating most factors are relevant at some point in time. Hence, fixing the inclusion probability to unity of any factor is unnecessary to conduct proper control. Moreover, including only the conventional factors leaves out much relevant information in the remainder of the asset pricing universe.

To examine the time variability in the importance of each factor further and its role in the difference across models’ abnormal returns, we depict in Figure 5 the time series of each factor’s inclusion probability. We focus in the remaining of the analysis mainly on the three-year event horizon to ease exposition. Conclusions are similar for the one-year and five-year horizons.

The inclusion probabilities generally show noticeable moves during NBER recessions. For instance, MOM, BAB, INF, and CONS spike in the two most recent recessions for the EW portfolio. It is clear from this figure in conjunction with Figure 2 and 4 that we can associate their patterns to
the disagreement between CTP-DMA and conventional models' abnormal returns. It is important, however, to note that the non-zero, yet non-unity inclusion probability of most factors plays a role in proper and efficient control, driving a further wedge between conventional models and CTP-DMA. To understand the difference between CTP-DMA and CTP-DMS better, we plot in Figure 6 the size of the single-best model at each time and in Figure 7 its composition of factors at each time point, which is to be compared with the average model size of the model averaging procedure in Figure 8 and the time series of inclusion probabilities.

We derive a number of interesting findings. First, the size of the DMS model is time-varying and fluctuates between four and nine factors (three and eleven) with an average of about seven (five) for the EW (VW) portfolio. The average model size using the entire asset pricing universe is generally increasing over time, from seven (six) to nine (seven), but with noticeable time-variation for the EW (VW) portfolio. This well exceeds that achievable from conventional models which includes at most five factors (and does so at all times). It, thus, appears increasingly relevant to allow for a broader set of factors to handle the dimensionality of expected returns.

Secondly, the composition of the best model varies over time, but the inclusion of certain factors are somewhat persistent. Except from ILLIQ and DFR all factors appear at some point in the best model for the EW portfolio whereas the VW portfolio requires use of all factors. Indeed, many of the variables not included in the conventional models are included in the best model. Moreover, HML, RMW, CMA, and MOM are not included at all times, pointing towards the difference between CTP-DMS and conventional models' abnormal returns. Additionally, none of the factors excluded from the DMS have zero inclusion probability on average. At times, their values are above 40% (cf. Figure 5), exceeding that of, among others, HML, MOM, and QMJ. Hence, treating them as globally irrelevant as in the CTP-DMS may be at risk of causing an omitted variable bias at those times, indicating why abnormal returns differ between CTP-DMS from CTP-DMA.

Thirdly, the size of the CTP-DMS is somewhat comparable to the average size of CTP-DMA though on average one (two) factor(s) smaller for the EW (VW) portfolio, cf. Table 3. Importantly, however, CTP-DMA exploits
information from the entire universe, weighting their relevance. CTP-DMS effectively puts all faith in the selected model at time $t$. The appropriateness of this can be inferred from the relative support for the best model at each time point as depicted (after integrating out the effect of $\delta_j$) in Figure 9.

The relative support for even the best model is low. It is higher than the equal share of $100/2^K = 0.003\%$ but still receives, at best, only a relative support of just above 3.5\%. This means that there is substantial cumulative relevant information in the remainder of the asset pricing universe besides that in the single best model. That is, if we were to use the most correct model from the model space in each period, one would put faith in models that are at best 96.5\% not the correct ones. For comparison, Figure 10 depicts the posterior model probability for each conventional model and the EW and VW portfolios.

The relative support for conventional models is generally higher than the equal share except for the conditional CAPM and an EW portfolio. Taking into account their low-dimensional specification, they actually perform relatively well in the asset pricing universe. However, the confidence in either model is never above 0.01\% and, as seen above, they all suffer from an omitted variable bias. Even though CTP-DMS uses the entire asset pricing universe better than those conventional models, it does so in a manner which is still at odds with the data. Rather, CTP-DMA utilizes information across the entire asset pricing universe in an efficient manner according to data to derive an accurate measure of abnormal returns.

It is relevant to compare these findings to Barillas and Shanken (2018). The average model size obtained by CTP-DMA is about seven-eight factors which compares to six in their preferred model. It should, however, be noted that their model is derived from an asset pricing universe that allows for at most six factors leaving it unclear whether other factors (such as those considered in this application) are also relevant. Moreover their preferred model effectively attaches a 100\% inclusion probability to all included factors at all times. Our results show that no single factor, not even from the conventional Fama-French set, is globally relevant at all times (at least on this sample). At the same time, the evidence for all factors is such that the dimensionality
of returns can be comprised into seven-eight factors from a set of 15 while still controlling for an omitted variable bias. This is interesting because this number exceeds what is sometimes found in statistical approaches, subsuming returns into 2-4 standard principal components (see, e.g., Kelly et al. (2019)). It goes, however, well in hand with the findings in Kozak et al. (2018b) that a few characteristic-based factors are insufficient to capture returns unless they are translated into orthogonal principal components. The reason is that most observed and traded factors are correlated as opposed to principal components, requiring the econometric technique to address a potential omitted variable bias. Our procedure indeed subsumes a large number of factors into a lower dimensional representation, and does so efficiently, whilst maintaining interpretability of the factors, which is otherwise lost by representing them as principal components as in Kozak et al. (2018b). It should also be noted that even though both Barillas and Shanken (2018) and our framework obtain posterior model probabilities, the use of those are very different. Consistent with an unconditional view, they select their preferred model at the end of the sample. We specify a conditional framework at the outset, enabling utilization of information contained at each time point recursively either via selection or averaging.

3.4.3. Empirical factor exposure

Figures 11 and 12 depict the estimated factor exposures using equation (20) for an EW and VW portfolio, respectively.

We readily observe that all factor exposures show some time varying behavior. For instance, the exposure to MKT (“market beta”) fluctuates around a mean of approximately 0.8 (1.0) for the EW (VW) portfolio. The factor exposures often show a noticeably move in periods with market stress accompanied by an increase in magnitude consistent with increasing risks during market stress. This is especially true for the exposure towards macroeconomic-based factors, MOM and BAB. Figure 13 depicts the posterior-probability weighted average of $\delta_t$.

It is apparent that coefficients vary differently over time, with estimated $\delta_t$ fluctuating around a mean of approximately 0.98. Especially periods surrounding crises show notable moves where periods leading up to crises seem
to be facing an increasing coefficient variation, which is then resolved gradually in the later parts of the crisis. Moreover, periods with large changes in the portfolio composition show notable moves which seem natural since the characteristics of the portfolio may change notably according to the inclusion and exclusion of different event firms. Interestingly, this may indicate that the firms entering the event portfolio possess quite different characteristics and factor exposures relative to those experiencing the event prior to this period, supporting the idea of a different signal in dividend initiations during the second part of our sample.

4. Application to characteristics-sorted portfolios

The analysis above has, at least, two important implications in a general asset pricing context. First, asset pricing models should contain quite a large number of factors (about 6-8) on average and exhibit substantial time-variation in the pertinence of various factors. Secondly, abnormal returns are strongly time-varying with DMA alphas sometimes negatively correlated with those from conventional models. To investigate whether these findings are general in a broader asset pricing context, we also apply our methodology to a large set of characteristics-sorted portfolios similarly in spirit to Giglio and Xiu (2018). We examine 202 portfolios over the 1980-2015 period comprised by 17 industry portfolios, 25 portfolios sorted by size and book-to-market ratio, 25 portfolios sorted by operating profitability and investment, 25 portfolios sorted by size and variance, 25 portfolios sorted by size and momentum, 35 portfolios sorted by size and net issuance, 25 portfolios sorted by size and accruals, and 25 portfolios sorted by size and stock market beta. This set of portfolios captures a vast cross-section of anomalies and exposures to various factors. As such, this application also serves as benchmark for the interpretation on the dividend initiation event effects documented above. Portfolio returns are obtained from Kenneth French’s library, and we employ identical model specifications as in the event study application. Since conclusions are similar for equal- and value-weighted portfolios, we comment only on the latter.

≪ Insert Table 5 about here ≫

Details on the portfolio constructions can be found on on Kenneth French’s website.
Table 5 summarizes the improvement of the DMA methodology relative to the DMS and conventional models as measured by logarithmic Bayes factors. The evidence in favour of the DMA relative to the DMS and conventional models are very strong on average, in the wording of Kass and Raftery (1995). In only a few cases some of the conventional models are preferred which is due to their smaller number of parameters relative to empirical fit, the Occam’s razor effect. This never happens for the conditional CAPM, 2/202 times for the conditional FF3, 3/202 times for the conditional FF4, and 4/202 times for the conditional FF5. The DMS is preferred over DMA only once. The improvements obtained from the event study above are generally in line with the average improvements obtained on this large cross-section of assets.

Tables 6-7 report summary statistics on, among other things, the number of factors and factor inclusion probabilities. On this basis, we may conclude that the results from the event study application carry over to the broad cross-section of stock returns in that the average model size is quite high (eight factors) with an average standard deviation of about one factor, showing some time-variation. This goes well in line with the findings in Feng et al. (2017) and Kozak et al. (2018b) that just a few characteristics-based factors, like in the FF3, cannot adequately capture variation in expected returns. Interestingly, we also note that there is generally little support for the dynamic, single model that best approximates the true factor structure. It receives an average of about 1% posterior model probability, yet at times reaches almost 100%. This best model is just above two factors smaller than DMA on average, however with a larger degree of time-variation. Lastly, we find strong evidence for inclusion of the size factor, SMB, relative to most of the other candidate factors. This goes well in line with the findings in Asness et al. (2018) of resurrecting the size effect, when controlling for the quality effect of the firms, achieved here by the allowed inclusion of the QMJ factor.

Table 8 reports a summary of the conditional abnormal returns and Table 9 reports their average correlation matrix across the various methods. According to this application, returns on the large panel of portfolios are generally considered less anomalous by our DMA methodology, relative to the benchmark methods typically considered in the literature. The time-variation of
alphas is substantial across all methods, yet larger in magnitude for our DMA comparing to the various Fama-French models, but similar to the CAPM. The alphas generally share a positive, yet semi-weak correlation. For comparison, the correlation between DMA and the FF3 abnormal returns in the dividend initiation application is in the range of 50-70%, but with notable negative correlation at times of market turmoil.

5. Concluding remarks

When evaluating the long-horizon impact of an event on firms’ performance, the validity of the conclusions depends crucially on accurate measurement of the normal return the event firms are expected to earn in the wake of the event. Using inaccurate measures may lead to wrong conclusions. We argue that such accurate measures must take into account potential misspecification of the asset pricing model while simultaneously allowing for shifting factor exposures over calendar-time. In response, we introduce a unified calendar-time portfolio framework which efficiently handles asset pricing model uncertainty and allows for time-varying alpha and factor exposures. The approach disciplines researchers’ use of asset pricing factors and assigns a probability measure to the appropriateness of (dynamically) selecting a single model that best approximates the true return factor structure or whether averaging across an asset pricing universe is desired.

We apply the proposed procedure to the widely debated hypothesis of long-horizon event effects following dividend initiations and resumptions in the 1980 to 2015 period. Our findings point to large economical and statistical differences in the results based on our unified framework and conventional models such as the conditional CAPM and versions of the Fama-French models. Choosing among conventional models has little influence on the conclusions and measurement of event effects, but they appear to be facing a severe omitted variable bias. Our finding of a generally negative event effect contrasts general conclusions in recent literature that event effects are largely positive, yet insignificant. An interesting route for further research is to test the robustness of existing findings on the long-horizon effect from other event studies such as equity offerings or mergers and acquisitions. Moreover, our findings from the large cross-section of portfolio returns suggest notable time variation in the dimensionality and composition of the factor structure in expected returns. Capturing this in structural asset pricing models is
interesting future research.
6. Acknowledgements

Left out in blinded manuscript.
References


7. Tables and figures

Table 1: Summary statistics of asset pricing factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Mean</th>
<th>Sd.</th>
<th>Max.</th>
<th>Min.</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>0.62</td>
<td>4.49</td>
<td>12.47</td>
<td>-23.22</td>
<td>1.09</td>
</tr>
<tr>
<td>SMB</td>
<td>0.12</td>
<td>2.95</td>
<td>19.18</td>
<td>-15.36</td>
<td>0.01</td>
</tr>
<tr>
<td>HML</td>
<td>0.26</td>
<td>3.02</td>
<td>13.91</td>
<td>-13.11</td>
<td>0.23</td>
</tr>
<tr>
<td>RMW</td>
<td>0.37</td>
<td>2.32</td>
<td>12.19</td>
<td>-17.67</td>
<td>0.32</td>
</tr>
<tr>
<td>CMA</td>
<td>0.30</td>
<td>2.03</td>
<td>9.51</td>
<td>-6.81</td>
<td>0.15</td>
</tr>
<tr>
<td>MOM</td>
<td>0.62</td>
<td>4.56</td>
<td>18.28</td>
<td>-34.58</td>
<td>0.74</td>
</tr>
<tr>
<td>BAB</td>
<td>0.95</td>
<td>3.61</td>
<td>12.89</td>
<td>-15.22</td>
<td>1.17</td>
</tr>
<tr>
<td>ILLIQ</td>
<td>0.30</td>
<td>5.80</td>
<td>8.67</td>
<td>-38.44</td>
<td>0.85</td>
</tr>
<tr>
<td>QMJ</td>
<td>0.49</td>
<td>2.53</td>
<td>12.89</td>
<td>-10.45</td>
<td>0.43</td>
</tr>
<tr>
<td>DFY</td>
<td>1.11</td>
<td>9.13</td>
<td>3.38</td>
<td>0.55</td>
<td>0.96</td>
</tr>
<tr>
<td>DFR</td>
<td>-0.01</td>
<td>1.55</td>
<td>7.37</td>
<td>-9.75</td>
<td>0.04</td>
</tr>
<tr>
<td>TMS</td>
<td>2.50</td>
<td>4.43</td>
<td>4.55</td>
<td>-3.65</td>
<td>2.47</td>
</tr>
<tr>
<td>INF</td>
<td>-0.19</td>
<td>1.19</td>
<td>4.04</td>
<td>-3.92</td>
<td>-0.22</td>
</tr>
<tr>
<td>IND</td>
<td>-0.98</td>
<td>4.72</td>
<td>17.66</td>
<td>-17.89</td>
<td>-0.84</td>
</tr>
<tr>
<td>CONS</td>
<td>0.40</td>
<td>3.23</td>
<td>12.81</td>
<td>-15.76</td>
<td>0.45</td>
</tr>
</tbody>
</table>

This table reports summary statistics for the set of candidate asset pricing factors as defined in Section 3.2 over the period of 1980 to 2015, covering a total of 432 monthly observations. Returns are in percentages.
Table 2: Logarithmic Bayes factors

<table>
<thead>
<tr>
<th>Comparison</th>
<th>$h = 1$ year</th>
<th></th>
<th></th>
<th>$h = 3$ years</th>
<th></th>
<th></th>
<th>$h = 5$ years</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EW</td>
<td>VW</td>
<td>EW</td>
<td>VW</td>
<td>EW</td>
<td>VW</td>
<td>EW</td>
<td>VW</td>
</tr>
<tr>
<td>CTP-DMA vs. CTP-DMS</td>
<td>0.38</td>
<td>19.03</td>
<td>3.08</td>
<td>18.73</td>
<td>5.31</td>
<td>14.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTP-DMA vs. CTP-CAPM</td>
<td>39.49</td>
<td>16.87</td>
<td>101.57</td>
<td>20.59</td>
<td>118.46</td>
<td>34.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTP-DMA vs. CTP-FF3</td>
<td>2.46</td>
<td>10.60</td>
<td>24.16</td>
<td>8.56</td>
<td>39.74</td>
<td>13.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTP-DMA vs. CTP-FF4</td>
<td>2.33</td>
<td>13.42</td>
<td>10.92</td>
<td>12.59</td>
<td>32.70</td>
<td>6.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTP-DMA vs. CTP-FF5</td>
<td>−1.22</td>
<td>11.78</td>
<td>16.64</td>
<td>4.77</td>
<td>24.43</td>
<td>7.36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table reports logarithmic Bayes factors of the comparison between the CTP-DMA methodology and the remaining methods under consideration with one-, three-, and five-year event horizons and both equally- and value-weighted (EW and VW, respectively) portfolios for the combined sample of dividend initiations and resumptions during the period 1983 to 2015. According to the rule-of-thumb interpretation rules of Kass and Raftery (1995), values below 1 are not worth more than a bare mention, values between 1-3 are positive in favour of the CTP-DMA, values between 3-5 are strongly in favour of the CTP-DMA, and values greater than 5 are very strongly in favour of the CTP-DMA.
Table 3: Long-horizon event portfolio abnormal returns

<table>
<thead>
<tr>
<th>Method</th>
<th>$h = 1$ year</th>
<th></th>
<th>$h = 3$ years</th>
<th></th>
<th>$h = 5$ years</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EW</td>
<td>VW</td>
<td>EW</td>
<td>VW</td>
<td>EW</td>
<td>VW</td>
</tr>
<tr>
<td>CTP-DMA</td>
<td>0.19***</td>
<td>-0.26</td>
<td>(7.42)</td>
<td>(7.35)</td>
<td>-0.09***</td>
<td>-0.45***</td>
</tr>
<tr>
<td>CTP-DMS</td>
<td>0.12***</td>
<td>-0.41**</td>
<td>(5.49)</td>
<td>(4.03)</td>
<td>0.29***</td>
<td>-0.55***</td>
</tr>
<tr>
<td>CTP-CAPM</td>
<td>0.64***</td>
<td>-0.08***</td>
<td>(1.00)</td>
<td>(1.00)</td>
<td>0.37***</td>
<td>-0.27***</td>
</tr>
<tr>
<td>CTP-FF3</td>
<td>0.70***</td>
<td>-0.00***</td>
<td>(3.00)</td>
<td>(3.00)</td>
<td>0.40***</td>
<td>-0.12***</td>
</tr>
<tr>
<td>CTP-FF4</td>
<td>0.66***</td>
<td>0.02***</td>
<td>(4.00)</td>
<td>(4.00)</td>
<td>0.49***</td>
<td>-0.10***</td>
</tr>
<tr>
<td>CTP-FF5</td>
<td>0.67***</td>
<td>0.00***</td>
<td>(5.00)</td>
<td>(5.00)</td>
<td>0.43***</td>
<td>-0.03***</td>
</tr>
</tbody>
</table>

This table reports average conditional alpha (in percentages) with one-, three-, and five-year event horizons and both equally- and value-weighted (EW and VW, respectively) portfolios for the combined sample of dividend initiations and resumptions during the period 1983 to 2015. For all methods, inference is based on conditional alpha over the full sample according to Section 2.2. Average model sizes are shown in parenthesis below each average conditional alpha. Superscripts ***, **, and * correspond to rejections at probability levels one, five, and ten percent, respectively.
Table 4: Summary of factor inclusion probabilities

<table>
<thead>
<tr>
<th>Factor</th>
<th>$h = 1$ year</th>
<th>$h = 3$ years</th>
<th>$h = 5$ years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EW</td>
<td>VW</td>
<td>EW</td>
</tr>
<tr>
<td>MKT</td>
<td>1.00</td>
<td>(0.00)</td>
<td>1.00</td>
</tr>
<tr>
<td>SMB</td>
<td>0.88</td>
<td>(0.00)</td>
<td>0.95</td>
</tr>
<tr>
<td>HML</td>
<td>0.63</td>
<td>(0.00)</td>
<td>0.67</td>
</tr>
<tr>
<td>RMW</td>
<td>0.48</td>
<td>(0.00)</td>
<td>0.52</td>
</tr>
<tr>
<td>CMA</td>
<td>0.46</td>
<td>(0.00)</td>
<td>0.47</td>
</tr>
<tr>
<td>MOM</td>
<td>0.47</td>
<td>(0.00)</td>
<td>0.57</td>
</tr>
<tr>
<td>BAB</td>
<td>0.43</td>
<td>(0.00)</td>
<td>0.41</td>
</tr>
<tr>
<td>ILLIQ</td>
<td>0.23</td>
<td>(0.00)</td>
<td>0.20</td>
</tr>
<tr>
<td>QMJ</td>
<td>0.49</td>
<td>(0.00)</td>
<td>0.57</td>
</tr>
<tr>
<td>DFY</td>
<td>0.43</td>
<td>(0.00)</td>
<td>0.51</td>
</tr>
<tr>
<td>DFR</td>
<td>0.35</td>
<td>(0.00)</td>
<td>0.32</td>
</tr>
<tr>
<td>TMS</td>
<td>0.30</td>
<td>(0.00)</td>
<td>0.39</td>
</tr>
<tr>
<td>INF</td>
<td>0.48</td>
<td>(0.00)</td>
<td>0.42</td>
</tr>
<tr>
<td>IND</td>
<td>0.22</td>
<td>(0.00)</td>
<td>0.23</td>
</tr>
<tr>
<td>CONS</td>
<td>0.57</td>
<td>(0.00)</td>
<td>0.66</td>
</tr>
</tbody>
</table>

This table reports the average and standard deviation (in parenthesis) across calendar-time of posterior inclusion probabilities of all candidate asset pricing factors across both equally- and value-weighted portfolios (EW and VW, respectively) with one-, three-, and five-year event horizons. For each time period and each asset pricing factor, the posterior inclusion probability is calculated as the sum of posterior model probabilities for all models where the asset pricing factor is included, see (18).
Table 5: Logarithmic Bayes factors

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Min.</th>
<th>Max.</th>
<th>Q1</th>
<th>Med.</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Equal-weighted portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMA vs. DMS</td>
<td>26.67</td>
<td>−8.41</td>
<td>77.92</td>
<td>17.38</td>
<td>21.02</td>
<td>32.99</td>
</tr>
<tr>
<td>DMA vs. CAPM</td>
<td>167.03</td>
<td>−5.98</td>
<td>353.91</td>
<td>107.77</td>
<td>177.70</td>
<td>227.24</td>
</tr>
<tr>
<td>DMA vs. FF3</td>
<td>65.67</td>
<td>−9.20</td>
<td>983.96</td>
<td>56.72</td>
<td>56.17</td>
<td>80.70</td>
</tr>
<tr>
<td>DMA vs. FF4</td>
<td>32.27</td>
<td>−3.34</td>
<td>275.21</td>
<td>10.52</td>
<td>32.14</td>
<td>45.69</td>
</tr>
<tr>
<td>DMA vs. FF5</td>
<td>52.11</td>
<td>1.46</td>
<td>266.29</td>
<td>21.08</td>
<td>51.13</td>
<td>66.55</td>
</tr>
<tr>
<td><strong>Panel B: Value-weighted portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMA vs. DMS</td>
<td>26.12</td>
<td>−14.96</td>
<td>73.93</td>
<td>18.90</td>
<td>24.82</td>
<td>33.46</td>
</tr>
<tr>
<td>DMA vs. CAPM</td>
<td>188.68</td>
<td>14.70</td>
<td>416.01</td>
<td>100.79</td>
<td>169.05</td>
<td>280.49</td>
</tr>
<tr>
<td>DMA vs. FF3</td>
<td>53.48</td>
<td>−0.92</td>
<td>332.17</td>
<td>19.87</td>
<td>34.32</td>
<td>65.78</td>
</tr>
<tr>
<td>DMA vs. FF4</td>
<td>34.82</td>
<td>−3.72</td>
<td>63.82</td>
<td>13.77</td>
<td>24.72</td>
<td>45.05</td>
</tr>
<tr>
<td>DMA vs. FF5</td>
<td>42.30</td>
<td>−2.78</td>
<td>328.57</td>
<td>10.73</td>
<td>23.11</td>
<td>51.67</td>
</tr>
</tbody>
</table>

This table reports summary statistics of the logarithmic Bayes factors of the comparison between the DMA methodology and the remaining methods under consideration applied to the 202 (equal- and value-weighted) characteristics-sorted portfolios during the period 1983-2015. According to the rule-of-thumb interpretation rules of Kass and Raftery (1995), values below 1 are not worth more than a bare mention, values between 1-3 are positive in favour of the DMA, values between 3-5 are strongly in favour of the DMA, and values greater than 5 are very strongly in favour of the DMA.
Table 6: Summary of empirical model uncertainty (equal-weighted portfolios)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Sd</th>
<th>Min.</th>
<th>Max.</th>
<th>Q1</th>
<th>Med.</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. factors (DMA)</td>
<td>8.23</td>
<td>0.90</td>
<td>2.32</td>
<td>12.06</td>
<td>7.44</td>
<td>8.30</td>
<td>9.08</td>
</tr>
<tr>
<td>No. factors (DMS)</td>
<td>6.00</td>
<td>1.69</td>
<td>1.00</td>
<td>13.00</td>
<td>5.00</td>
<td>6.00</td>
<td>7.00</td>
</tr>
<tr>
<td>Highest post. prob. (%)</td>
<td>1.18</td>
<td>1.15</td>
<td>0.02</td>
<td>98.88</td>
<td>0.19</td>
<td>0.37</td>
<td>0.73</td>
</tr>
<tr>
<td>$\delta_t$ (100)</td>
<td>97.98</td>
<td>0.34</td>
<td>97.00</td>
<td>98.89</td>
<td>97.72</td>
<td>98.06</td>
<td>98.27</td>
</tr>
<tr>
<td>MKT</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SMB</td>
<td>0.81</td>
<td>0.11</td>
<td>0.00</td>
<td>1.00</td>
<td>0.59</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>HML</td>
<td>0.58</td>
<td>0.17</td>
<td>0.00</td>
<td>1.00</td>
<td>0.40</td>
<td>0.51</td>
<td>0.79</td>
</tr>
<tr>
<td>RMW</td>
<td>0.49</td>
<td>0.14</td>
<td>0.00</td>
<td>1.00</td>
<td>0.39</td>
<td>0.45</td>
<td>0.56</td>
</tr>
<tr>
<td>CMA</td>
<td>0.48</td>
<td>0.11</td>
<td>0.00</td>
<td>1.00</td>
<td>0.37</td>
<td>0.44</td>
<td>0.55</td>
</tr>
<tr>
<td>MOM</td>
<td>0.60</td>
<td>0.24</td>
<td>0.00</td>
<td>1.00</td>
<td>0.38</td>
<td>0.55</td>
<td>0.89</td>
</tr>
<tr>
<td>BAB</td>
<td>0.55</td>
<td>0.13</td>
<td>0.00</td>
<td>1.00</td>
<td>0.38</td>
<td>0.49</td>
<td>0.74</td>
</tr>
<tr>
<td>ILLIQ</td>
<td>0.38</td>
<td>0.16</td>
<td>0.00</td>
<td>1.00</td>
<td>0.31</td>
<td>0.38</td>
<td>0.45</td>
</tr>
<tr>
<td>QMJ</td>
<td>0.53</td>
<td>0.15</td>
<td>0.00</td>
<td>1.00</td>
<td>0.39</td>
<td>0.47</td>
<td>0.65</td>
</tr>
<tr>
<td>DFY</td>
<td>0.13</td>
<td>0.08</td>
<td>0.01</td>
<td>1.00</td>
<td>0.39</td>
<td>0.43</td>
<td>0.47</td>
</tr>
<tr>
<td>DFR</td>
<td>0.41</td>
<td>0.10</td>
<td>0.00</td>
<td>1.00</td>
<td>0.34</td>
<td>0.41</td>
<td>0.46</td>
</tr>
<tr>
<td>TMS</td>
<td>0.41</td>
<td>0.08</td>
<td>0.00</td>
<td>1.00</td>
<td>0.36</td>
<td>0.40</td>
<td>0.45</td>
</tr>
<tr>
<td>INF</td>
<td>0.53</td>
<td>0.13</td>
<td>0.01</td>
<td>1.00</td>
<td>0.41</td>
<td>0.48</td>
<td>0.63</td>
</tr>
<tr>
<td>IND</td>
<td>0.46</td>
<td>0.14</td>
<td>0.00</td>
<td>1.00</td>
<td>0.36</td>
<td>0.44</td>
<td>0.53</td>
</tr>
<tr>
<td>CONS</td>
<td>0.55</td>
<td>0.15</td>
<td>0.00</td>
<td>1.00</td>
<td>0.40</td>
<td>0.48</td>
<td>0.69</td>
</tr>
</tbody>
</table>

This table reports summary statistics of various metrics associated with empirical asset pricing model uncertainty as estimated by the DMA and DMS methodologies applied to the 202 equal-weighted characteristics-sorted portfolios during the period 1983-2015.
Table 7: Summary of empirical model uncertainty (value-weighted portfolios)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Sd.</th>
<th>Min.</th>
<th>Max.</th>
<th>Q1</th>
<th>Med.</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. factors (DMA)</td>
<td>8.13</td>
<td>0.79</td>
<td>1.59</td>
<td>12.13</td>
<td>7.60</td>
<td>8.23</td>
<td>8.80</td>
</tr>
<tr>
<td>No. factors (DMS)</td>
<td>5.97</td>
<td>1.57</td>
<td>1.00</td>
<td>14.00</td>
<td>5.00</td>
<td>6.00</td>
<td>7.00</td>
</tr>
<tr>
<td>Highest post. prob. (%)</td>
<td>1.05</td>
<td>1.23</td>
<td>0.02</td>
<td>98.52</td>
<td>0.17</td>
<td>0.32</td>
<td>0.66</td>
</tr>
<tr>
<td>(\delta_t \cdot (100))</td>
<td>98.06</td>
<td>0.30</td>
<td>97.00</td>
<td>98.96</td>
<td>97.86</td>
<td>98.11</td>
<td>98.30</td>
</tr>
<tr>
<td>MKT</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SMB</td>
<td>0.80</td>
<td>0.11</td>
<td>0.00</td>
<td>1.00</td>
<td>0.58</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>HML</td>
<td>0.58</td>
<td>0.17</td>
<td>0.00</td>
<td>1.00</td>
<td>0.40</td>
<td>0.51</td>
<td>0.80</td>
</tr>
<tr>
<td>RMW</td>
<td>0.50</td>
<td>0.14</td>
<td>0.00</td>
<td>1.00</td>
<td>0.39</td>
<td>0.46</td>
<td>0.58</td>
</tr>
<tr>
<td>CMA</td>
<td>0.50</td>
<td>0.10</td>
<td>0.00</td>
<td>1.00</td>
<td>0.37</td>
<td>0.45</td>
<td>0.61</td>
</tr>
<tr>
<td>MOM</td>
<td>0.51</td>
<td>0.16</td>
<td>0.00</td>
<td>1.00</td>
<td>0.35</td>
<td>0.44</td>
<td>0.64</td>
</tr>
<tr>
<td>BAB</td>
<td>0.49</td>
<td>0.10</td>
<td>0.00</td>
<td>1.00</td>
<td>0.36</td>
<td>0.43</td>
<td>0.60</td>
</tr>
<tr>
<td>ILLIQ</td>
<td>0.39</td>
<td>0.12</td>
<td>0.00</td>
<td>1.00</td>
<td>0.31</td>
<td>0.38</td>
<td>0.45</td>
</tr>
<tr>
<td>QMJ</td>
<td>0.51</td>
<td>0.14</td>
<td>0.00</td>
<td>1.00</td>
<td>0.40</td>
<td>0.46</td>
<td>0.61</td>
</tr>
<tr>
<td>DFY</td>
<td>0.43</td>
<td>0.07</td>
<td>0.01</td>
<td>1.00</td>
<td>0.40</td>
<td>0.43</td>
<td>0.47</td>
</tr>
<tr>
<td>DFR</td>
<td>0.41</td>
<td>0.12</td>
<td>0.01</td>
<td>1.00</td>
<td>0.34</td>
<td>0.41</td>
<td>0.47</td>
</tr>
<tr>
<td>TMS</td>
<td>0.46</td>
<td>0.08</td>
<td>0.00</td>
<td>0.99</td>
<td>0.36</td>
<td>0.40</td>
<td>0.44</td>
</tr>
<tr>
<td>INF</td>
<td>0.51</td>
<td>0.14</td>
<td>0.00</td>
<td>1.00</td>
<td>0.42</td>
<td>0.50</td>
<td>0.66</td>
</tr>
<tr>
<td>IND</td>
<td>0.47</td>
<td>0.14</td>
<td>0.00</td>
<td>1.00</td>
<td>0.36</td>
<td>0.44</td>
<td>0.55</td>
</tr>
<tr>
<td>CONS</td>
<td>0.57</td>
<td>0.15</td>
<td>0.00</td>
<td>1.00</td>
<td>0.41</td>
<td>0.52</td>
<td>0.72</td>
</tr>
</tbody>
</table>

This table reports summary statistics of various metrics associated with empirical asset pricing model uncertainty as estimated by the DMA and DMS methodologies applied to the 202 value-weighted characteristics-sorted portfolios during the period 1983-2015.
<table>
<thead>
<tr>
<th></th>
<th>DMA</th>
<th>DMS</th>
<th>CAPM</th>
<th>FF3</th>
<th>FF4</th>
<th>FF5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Equal-weighted portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean $</td>
<td>\alpha_t</td>
<td>$</td>
<td>0.39</td>
<td>0.49</td>
<td>0.45</td>
<td>0.33</td>
</tr>
<tr>
<td>Sd. $</td>
<td>\alpha_t</td>
<td>$</td>
<td>0.44</td>
<td>0.62</td>
<td>0.43</td>
<td>0.27</td>
</tr>
<tr>
<td>Max. $</td>
<td>\alpha_t</td>
<td>$</td>
<td>6.22</td>
<td>9.17</td>
<td>4.43</td>
<td>2.28</td>
</tr>
<tr>
<td>Min. $</td>
<td>\alpha_t</td>
<td>$</td>
<td>-5.63</td>
<td>-5.47</td>
<td>-3.49</td>
<td>-2.88</td>
</tr>
<tr>
<td>Q1 $</td>
<td>\alpha_t</td>
<td>$</td>
<td>0.11</td>
<td>0.13</td>
<td>0.16</td>
<td>0.11</td>
</tr>
<tr>
<td>Med. $</td>
<td>\alpha_t</td>
<td>$</td>
<td>0.26</td>
<td>0.29</td>
<td>0.33</td>
<td>0.24</td>
</tr>
<tr>
<td>Q3 $</td>
<td>\alpha_t</td>
<td>$</td>
<td>0.50</td>
<td>0.60</td>
<td>0.62</td>
<td>0.43</td>
</tr>
<tr>
<td>No. 1% rej.</td>
<td>0.73</td>
<td>0.91</td>
<td>0.83</td>
<td>0.71</td>
<td>0.88</td>
<td>0.84</td>
</tr>
<tr>
<td>No. 5% rej.</td>
<td>0.81</td>
<td>0.92</td>
<td>0.86</td>
<td>0.77</td>
<td>0.90</td>
<td>0.89</td>
</tr>
<tr>
<td>No. 10% rej.</td>
<td>0.82</td>
<td>0.94</td>
<td>0.88</td>
<td>0.80</td>
<td>0.92</td>
<td>0.91</td>
</tr>
<tr>
<td><strong>Panel B: Value-weighted portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean $</td>
<td>\alpha_t</td>
<td>$</td>
<td>0.31</td>
<td>0.36</td>
<td>0.36</td>
<td>0.26</td>
</tr>
<tr>
<td>Sd. $</td>
<td>\alpha_t</td>
<td>$</td>
<td>0.34</td>
<td>0.48</td>
<td>0.37</td>
<td>0.25</td>
</tr>
<tr>
<td>Max. $</td>
<td>\alpha_t</td>
<td>$</td>
<td>6.18</td>
<td>6.39</td>
<td>4.39</td>
<td>1.96</td>
</tr>
<tr>
<td>Min. $</td>
<td>\alpha_t</td>
<td>$</td>
<td>-4.24</td>
<td>-5.27</td>
<td>-4.25</td>
<td>-2.96</td>
</tr>
<tr>
<td>Q1 $</td>
<td>\alpha_t</td>
<td>$</td>
<td>0.50</td>
<td>0.11</td>
<td>0.12</td>
<td>0.09</td>
</tr>
<tr>
<td>Med. $</td>
<td>\alpha_t</td>
<td>$</td>
<td>0.22</td>
<td>0.24</td>
<td>0.26</td>
<td>0.19</td>
</tr>
<tr>
<td>Q3 $</td>
<td>\alpha_t</td>
<td>$</td>
<td>0.41</td>
<td>0.49</td>
<td>0.49</td>
<td>0.35</td>
</tr>
<tr>
<td>No. 1% rej.</td>
<td>0.61</td>
<td>0.90</td>
<td>0.68</td>
<td>0.58</td>
<td>0.73</td>
<td>0.66</td>
</tr>
<tr>
<td>No. 5% rej.</td>
<td>0.66</td>
<td>0.92</td>
<td>0.78</td>
<td>0.66</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>No. 10% rej.</td>
<td>0.70</td>
<td>0.93</td>
<td>0.82</td>
<td>0.71</td>
<td>0.80</td>
<td>0.81</td>
</tr>
</tbody>
</table>

This table reports summary statistics of conditional abnormal returns $\alpha_t$, including the number of times the probability of zero alpha is less than 1%, 5%, and 10% levels using Section 2.2, as applied to the 202 (equal- and value-weighted) characteristics-sorted portfolios during the period 1983-2015.
Table 9: Correlation matrix of abnormal returns

<table>
<thead>
<tr>
<th></th>
<th>DMA</th>
<th>DMS</th>
<th>CAPM</th>
<th>FF3</th>
<th>FF4</th>
<th>FF5</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMA</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DMS</td>
<td>0.82</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CAPM</td>
<td>0.35</td>
<td>0.27</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FF3</td>
<td>0.40</td>
<td>0.30</td>
<td>0.66</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FF4</td>
<td>0.48</td>
<td>0.38</td>
<td>0.54</td>
<td>0.81</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>FF5</td>
<td>0.49</td>
<td>0.39</td>
<td>0.50</td>
<td>0.80</td>
<td>0.78</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Panel A: Equal-weighted portfolios

Panel B: Value-weighted portfolios

This table reports the average correlation matrix between the conditional abnormal returns among the models under consideration, as applied to the 202 (equal- and value-weighted) characteristics-sorted portfolios during the period 1983-2015.
This figure depicts the distribution of events (dividends and resumptions) in each year from 1980 to 2010. The method of identification is described in Section 3.1. Red bars indicate years where (part of) an NBER recession occurred.
This figure depicts the abnormal returns (alpha) for the sample period 1983 to 2013 using the CTP-DMA procedure and conventional models for an event horizon equal to three years and an equally-weighted portfolio. Returns are in percentages. The upper figure depicts the CTP-DMA alpha in black solid line where the red (light red) shaded area indicates the 68% (95%) posterior credible intervals, based on $\hat{\Sigma}_t$. The lower figure includes in dot-dashed lines alpha obtained from the conventional models; conditional CTP-CAPM (red), CTP-FF3 (blue), CTP-FF4 (orange), and CTP-FF5 (green). Dashed vertical lines (shaded areas) indicate NBER recessions in the upper (lower) figure.
Figure 3: Conditional abnormal returns (value-weighted, \( h = 1 \))

This figure depicts the abnormal returns (alpha) for the sample period 1983 to 2011 using the CTP-DMA procedure for an event horizon equal to one year and a value-weighted portfolio. Returns are in percentages. The upper figure depicts the CTP-DMA alpha in black solid line where the red (light red) shaded area indicates the 68% (95%) posterior credible intervals, based on \( \hat{\Sigma}_t \). The lower figure depicts the probability of the null hypothesis of zero alpha using (21) and (22). Dashed vertical lines (shaded areas) indicate NBER recessions in the upper (lower) figure.
This figure depicts the abnormal returns (alpha) for the sample period 1983 to 2013 using the CTP-DMA procedure and conventional models for an event horizon equal to three years and a value-weighted portfolio. Returns are in percentages. The upper figure depicts the CTP-DMA alpha in black solid line where the red (light red) shaded area indicates the 68% (95%) posterior credible intervals, based on $\hat{\Sigma}_t$. The lower figure includes in dot-dashed lines alpha obtained from the conventional models; conditional CTP-CAPM (red), CTP-FF3 (blue), CTP-FF4 (orange), and CTP-FF5 (green). Dashed vertical lines (shaded areas) indicate NBER recessions in the upper (lower) figure.
Figure 5: Time-varying inclusion probabilities

(a) MKT

(b) SMB

(c) HML

(d) RMW

(e) CMA

(f) MOM

(g) BAB

(h) ILLIQ
This figure consists of 15 graphs, each depicting the posterior inclusion probability of each candidate asset pricing factor for the 1983 to 2013 period with an equally-weighted (black line) and value-weighted (red line) portfolio and three-year event horizon. At each time point, the posterior inclusion probability for each candidate asset pricing factor is calculated as the sum of posterior model probabilities for all models where the asset pricing factor is included. Shaded areas indicate NBER recessions.
Figure 6: Size of “most correct” model

This figure depicts the number of included factors (size) of the model with the highest posterior model probability in the 1983 to 2013 period using an equally-weighted (black line) and value-weighted (red line) portfolio and three-year event horizon. Shaded areas indicate NBER recessions.
Figure 7: Factor composition of “most correct” model.

This figure depicts the factor composition of the model with the highest posterior model probability at each time point in the 1983 to 2013 period using an equally-weighted (upper figure) and value-weighted (lower figure) portfolio and three-year event horizon. A coloured circle indicates inclusion in the dynamically selected model. Shaded areas indicate NBER recessions.
Figure 8: Average size of candidate factor models

This figure depicts the average number of included factors (size) across the entire pool of candidate models, as given by (17), in the 1983 to 2013 period, using an equally-weighted (black line) and value-weighted (red line) portfolio and three-year event horizon. Shaded areas indicate NBER recessions.
Figure 9: Posterior model probability of “most correct” model

This figure depicts the highest posterior model probability (in percentage) in each time point across the entire model space in the 1983 to 2013 period using an equally-weighted (black line) and value-weighted (red line) portfolio and three-year event horizon. Shaded areas indicate NBER recessions.
Figure 10: Posterior model probabilities of conventional models

This figure depicts the posterior model probability (in percentage) in each time period for each of the conventional models: conditional CTP-CAPM (red), CTP-FF3 (blue), CTP-F4 (orange), and CTP-FF5 (green) for an equally-weighted (upper figure) and value-weighted (lower figure) portfolio and three-year event horizon. The horizontal, solid black line indicates the equal share of $1/2^K$. The curves are slightly smoothened to promote interpretation. Shaded areas indicate NBER recessions.
Figure 11: Time-varying factor exposures (equally-weighted, \( h = 3 \))

(a) MKT

(b) SMB

(c) HML

(d) RMW

(e) CMA

(f) MOM

(g) \( \beta_A \)

(h) ILLIQ
Figure 11 (Cont.): Time-varying factor exposures (equally-weighted, \( h = 3 \))

This figure consists of 15 graphs, each depicting the estimated exposures (betas) to each candidate asset pricing factor for the 1983 to 2013 period for an equally-weighted portfolio and three-year event horizon. At each time point the estimates are computed as in equation (20). The red (light red) shaded area indicates the 68% (95%) posterior credible intervals, based on \( \hat{\Sigma}_t \). Dashed vertical lines indicate NBER recessions.
Figure 12: Time-varying factor exposures (value-weighted, $h = 3$)
Figure 12 (Cont.): Time-varying factor exposures (value-weighted, $h = 3$)

This figure consists of 15 graphs, each depicting the estimated exposures (betas) to of each candidate asset pricing factor for the 1983 to 2013 period for a value-weighted portfolio and three-year event horizon. At each time point the estimates are computed as in equation (20). The red (light red) shaded area indicates the 68% (95%) posterior credible intervals, based on $\hat{\Sigma}_t$. Dashed vertical lines indicate NBER recessions.
This figure depicts the posterior-probability weighted average of the time variation in coefficients in the 1983 to 2013 period for an equally-weighted (black line) and value-weighted (red line) portfolio and three-year event horizon. Shaded areas indicate NBER recessions.