Econometric Theory, **38**, 2022, 1175–1220.
doi:10.1017/S0266466622000020

SPECTRAL FINANCIAL ECONOMETRICS

FEDERICO M. BANDI
Carey Business School, Johns Hopkins University

ANDREA TAMONI
Rutgers Business School, Rutgers University

We survey the literature on spectral regression estimation. We present a cohesive framework designed to model dependence on frequency in the response of economic time series to changes in the explanatory variables. Our emphasis is on the statistical structure and on the economic interpretation of time-domain specifications needed to obtain horizon effects over *frequencies*, over *scales*, or upon *aggregation*. To this end, we articulate our discussion around the role played by lead-lag effects in the explanatory variables as drivers of differential information across horizons. We provide perspectives for future work throughout.

1. INTRODUCTION

In macroeconomics, the idea of studying relations between variables over different time horizons dates back to, at least, the work of Burns and Mitchell (1946) on economic cycles. In their empirical account of income dynamics across occupations, Friedman and Kuznets (1954) imply a decomposition of the income process into—essentially—a permanent and a transitory component. Famously, the same decomposition is central to the study of consumption in Friedman (1957). Following these and other early contributions, it is now rather standard practice to examine economic relations by focusing on a particular set of *frequencies*, *scales*, or *horizons*, the two most common time frames being the business cycle, generally defined as fluctuations between (roughly) 8 and 32 quarters (see, e.g., Englund, Persson, and Svensson, 1992; King and Watson, 1996), and the medium/long run (see, e.g., Comin and Gertler, 2006, who study horizons up to 50 years). An analogous taxonomy is used in macro-finance.

A natural inferential tool to evaluate economic models in which the relation between variables depends on frequency is spectrum regression. As is well known, spectrum regression effectively amounts to a two-step procedure. First, dependent and exogenous variables are Fourier-transformed to the frequency domain. Second,

We dedicate this article to Peter C.B. Phillips on the occasion of the conference "A Celebration of Peter Phillips' 40 Years at Yale," New Haven, October 19–20, 2018. We thank Peter and conference participants for their useful comments. We are especially grateful to the Editor, Guido Kuersteiner, and two anonymous reviewers for their careful reading and many helpful suggestions. Address correspondence to Federico M. Bandi, Carey Business School, Johns Hopkins University, Baltimore, MD, USA; e-mail: fbandi1@jhu.edu.

© The Author(s), 2022. Published by Cambridge University Press.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

the parameter estimates are obtained from the transformed data by employing all fundamental frequencies or a subset of them deemed of economic interest (c.f. Hannan, 1963a, 1963b). Alternative (other than Fourier) transforms yielding time-scale representations, like the Haar wavelets (Haar, 1910), as well as more traditional temporal aggregation methods, also provide helpful information about horizon-specific dependencies.

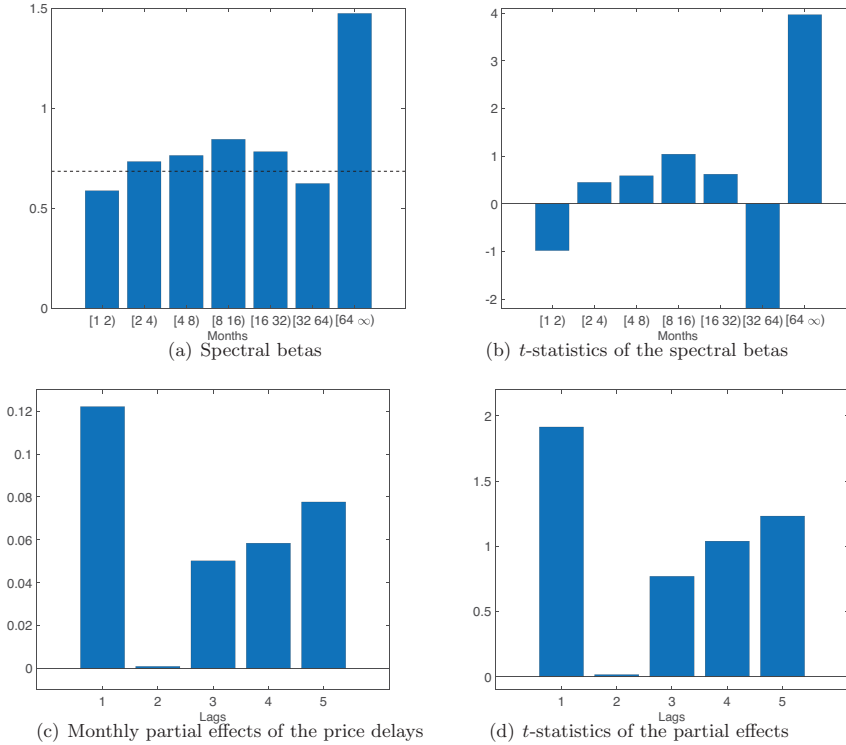
If the relation between economic variables changes across frequencies, scales, or horizons, a possibly nontrivial *lead-lag* structure should exist between these variables in their time-domain specification. While this observation is rather intuitive (and implicit in existing theoretical work in spectral econometrics), it has—arguably—not led to complete explorations into, for example, the economic interpretation of dynamic dependence structures in the time domain and their implications. Yet, stimulating recent work, for example, in finance, points to the importance of short-, medium-, and long-term delayed adjustments in the determination of time series of interest given predictors (c.f. Section 2.3).

In light of this observation, we organize this article around the mapping between heterogeneity over frequencies (as captured by spectrum regression), scales (as given by multiresolution analysis and more recent approaches), and horizons (as obtained by aggregation) and the resulting lead-lag structure in the time-domain specification of the model. We hope that, by making salient the link between variation over frequencies in the frequency/scale domain and asynchronous adjustments to information in the time domain, this article will aid economic interpretation of inferential procedures which may begin, as sometimes is the case, with frequency-domain representations and related methods of inference. More generally, we also hope that it will foster a broader use of spectral methods—broadly defined here as the set of inferential methods revealing of frequency-specific dynamics, whether explicitly in the frequency domain or not—as a way to achieve both more effective identification of rich time-domain structures and superior economic applicability.

Let us illustrate by way of an example. The use of spectral methods to study the nature of economic relations across horizons is arguably less common in finance than in other areas of economics. This is somewhat surprising. The most fundamental measure of (systematic) risk in financial markets is market beta (Sharpe, 1964; Lintner, 1965), the regression slope between excess returns on an asset and excess returns on the (mean-variance efficient, Markowitz, 1952) market portfolio, a portfolio generally proxied by a market index (c.f. Roll, 1977, for a well-known critique). In the data, this classical quantity varies across horizons, a result which has often been documented using returns computed over alternative holding periods, that is, by aggregation. We expand on this observation in Sections 2.2 and 2.3.

In Figure 1, we, instead, employ the spectral (time-scale) regression approach recently put forward by Bandi et al. (2021) and reviewed in Section 4 (to which we refer the reader for more details). We focus on one of the two stocks used by Gilbert et al. (2014) in their illustration of horizon effects (through aggregation)

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press



(a) Spectral betas (b) t-statistics of the spectral betas

(c) Monthly partial effects of the price delays (d) t-statistics of the partial effects

FIGURE 1. Berkshire Hathaway: seven spectral betas and partial effects of five delayed adjustments. The top panels report seven estimated spectral betas for BRK (corresponding to economic fluctuations over time horizons specified on the horizontal axis) and Newey-West-adjusted t-statistics. The dotted line in Panel (a) is the standard beta (0.69). The standard beta is a weighted average of the spectral betas with weights given by the relative variances of the corresponding market components (i.e., the scale-specific market variances relative to overall market variance; c.f. Bandi et al., 2021). The t-statistics are centered at the standard beta. The bottom panels report the estimated delayed partial effects from regressing BRK returns on the market return and five lags of the market return and the respective Newey-West-adjusted t-statistics. The market return is value-weighted. The data is monthly and the sample period is November 1976 to December 2017.

in the computation of market betas, namely, Berkshire Hathaway (BRK). Panel (a) reports seven *spectral betas*. The first six are the betas between components of the value-weighted market return process and the return process of BRK corresponding to economic fluctuations spanning dyadic horizons between 2^{j-1} months and 2^j months, with $j = 1, 2, \ldots, 6$. The last beta refers to the components spanning the residual horizon: 2^6 months to infinity.

We note that the classical beta, reported in Panel (a) by way of a dotted line, is simply a weighted average of the spectral betas with weights given by the variance of the corresponding market component relative to the overall variance of the

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

market return (Bandi et al., 2021).[1] This beta is significantly below unity with a t-statistic (relative to 1) of -3.18.

Panel (a) shows that the amount of systematic risk of BRK—as proxied by the spectral betas—increases, albeit not monotonically, as a function of the time horizon, and is particularly large in correspondence with the lowest frequency. Panel (b) reports Newey-West-adjusted t-statistics centered with respect to a natural benchmark, the classical beta (corresponding, again, to the dotted line in Panel (a)). The figure shows that the last two spectral betas, the last one in particular (t-statistic of 3.96), are statistically considerably different from what would be estimated when ignoring horizon effects.[2] We also note that if we were to report individual t-statistics for a classical zero null instead, they would all be larger than 5 with the sole exception of the penultimate one, which would be larger than 3.

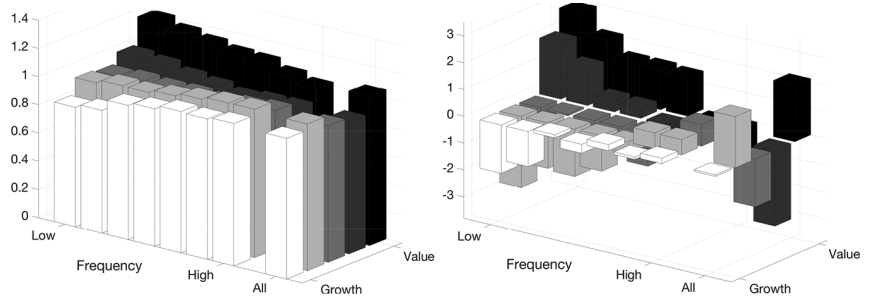
Figure 2 provides, again, seven spectral betas and, in column "All," market betas (Panel (a)) along with Newey-West-adjusted t-statistics centered, once more, around suitable, restricted betas (Panel (b)) for five value-sorted quintile portfolios. The overall market betas for the five portfolios are close to 1. For these betas, we therefore report t-statistics in deviation from unity. Quintiles 1–3 have t-statistics equal to 0.06, 1.90, and -1.78. Thus, these portfolios have market betas that are statistically not different from unity. The fourth and the fifth portfolio, on the other hand, have betas that differ from unity with t-statistics equal to -2.81 and 2.11, respectively. We observe that the spectral betas decline over frequency for the first two quintiles but increase over frequency for the last two quintiles, both the decrease and the increase being more evident for the more extreme (first and fifth) quintiles. Because high book-to-market (value) stocks have spectral betas which increase with the horizon, the increase in the spectral betas of BRK is consistent—based on this metric, at least—with several official corporate statements in Warren Buffett's yearly letter to shareholders stating or implying that Berkshire's strategy amounts to a long-term, buy-and-hold, "value" investment approach. As earlier in the BRK case, the benchmark used in the computation of the centered t-statistics associated with the spectral betas is the traditional market beta (rather than unity), c.f. Panel (b). Up to scale 5, no spectral beta is significantly different from the market beta. At scale 6, the spectral beta for the fifth (value) portfolio is statistically larger than the market beta with a t-statistic of 2.09. Moving to the lowest frequency, we document statistically significant deviations from the market betas for all portfolios except for the middle one, the t-statistics being -2.32, -2.65, -1.04, 2.09, and 2.98.[3] In sum, the market beta on the growth portfolio is not different from unity, but its lowest frequency-specific counterpart

[1] The spectral betas in Figure 1 are computed using a nonparametric Haar filter (Section 4.1). The mapping between the Wold representation-based spectral betas of Bandi et al. (2021) and the Haar-based spectral betas reported here is discussed in the Appendix of Bandi and Tamoni (2021).

[2] A formal $F-$type test of the hypothesis that the spectral betas are the same rejects at the 1% level (with the rejection being driven by the two lowest frequencies).

[3] Consistent with the BRK case, the unreported (less informative) t-statistics for a zero null would all be extremely large.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press



(a) Spectral and standard betas (b) t-statistics of the spectral and standard betas

FIGURE 2. Value-sorted portfolios: spectral and standard betas. The left panel reports seven estimated spectral betas (corresponding to economic fluctuations over the time horizons specified in Panel (a) of Figure 1) for five quintiles of value-sorted portfolios along with standard betas (in column "All"). The right panel reports the corresponding Newey-West-adjusted t-statistics. The standard betas are weighted averages of the corresponding spectral betas with weights given by the relative variances of the corresponding market components (i.e., scale-specific market variances relative to the overall market variance; c.f. Bandi et al., 2021). The t-statistics for the market betas in "All" are centered at 1. Those for the spectral betas are centered at the market betas. The market return is value-weighted. The data is monthly and the sample period is November 1976 to December 2017.

is significantly below 1. As for the value portfolio, the market beta is larger than unity and its lowest frequency-specific counterpart is larger than the market beta.

Turning now to a distributed lag model for BRK, we run a regression of BRK returns on the current market return as well as five monthly lags of the market return. The coefficients associated with past values of the market returns are reported in Panel (c) of Figure 1. Their t-statistics (for a zero null) are in Panel (d). The first delayed adjustment is the largest, both in numerical terms and in terms of statistical significance. Its positive sign finds information-based justifications in the existing finance literature. Value investing is viewed as not being glamorous and may, therefore, be either less understood and "opaque" (Gilbert et al., 2014) or simply neglected (Hou and Moskowitz, 2005). In both cases, the result is a delayed adjustment to market updates.

Importantly, we will show in Section 4 that a time-series model with one delayed adjustment (i.e., a model inclusive of the only adjustment which is significant at the 10% level, c.f. Panel (d) of Figure 1) is rather effective in capturing the increasing, albeit at a decreasing rate, first four spectral betas in Panel (a) of Figure 1. The richer structure of the spectral betas at lower frequencies—and, in particular, the lowest frequency beta—requires longer time-domain delays which are statistically harder to identify but are likely to be, economically, as consequential as the short-term delays described earlier. The economic importance of hard-to-detect long delays is, of course, not specific to the recent finance literature (see, e.g., Fuster, Laibson, and Mendel, 2010, and the references therein). In finance, it is key to a complete understanding of low-frequency risk dynamics and their pricing.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Why should asset returns adjust to factors with lags, thereby contradicting a central tenet of some fundamental asset pricing models? We do not provide an answer in this article other than discussing contributions which do offer some perspectives. Instead, we focus on reviewing methods of statistical inference which will raise more questions—laid out throughout—but have the potential to constitute an interesting arsenal for informative explorations.

We proceed as follows. Section 2 contains a nontechnical description of (a selection of) theoretical and applied contributions in spectral (regression) analysis. Selected works in macroeconomics and finance are discussed in Sections 2.1 and 2.2, respectively. Section 2.3 further positions the focus of this article within recent contributions in the finance literature. Sections 3–5 provide a more technical framework focusing on the logical connections between methods of linear inference in the frequency domain, in the time-scale domain, and through aggregation. In all cases, we comment on the mapping with the corresponding time-domain representations. Section 6 concludes by providing further directions for future investigations.

2. THE ECONOMETRIC LITERATURE

The literature on the spectral analysis of time series is vast (see, e.g., the review chapter of Granger and Watson, 1984, for a coverage of early themes).[4] We do not attempt to review it in its entirety here but, instead, focus on a specific subject with interesting theoretical implications and far-reaching empirical applicability: spectrum regression.

Spectrum regression with excluded frequencies (band spectrum regression) is set forth in Hannan (1963a, 1963b). In its original formulation, the exclusion of a band of (high) frequencies is suggested as a way to reduce measurement error in the explanatory variables in linear regression, an "open" model in the jargon of Robinson (1977). The usefulness of frequency elimination is also discussed in Hannan (Hannan, 1969, p. 584) for the purpose of the estimation of a linear discrete-time model, a type of "closed" model, in which a specific frequency band is affected by noise. A thorough treatment in the context of "closed" models in continuous time as well as in discrete time is provided by Robinson (1977) who recommends the omission of suitable, small clusters of frequencies to reduce seasonal noise or measurement error (see, also, Thomson, 1986, for a related approach). A Chow-type F-test of parameter stability across frequency bands is proposed in Engle (1974) and further explored in Engle (1978). The usefulness of the test hinges on the fact that, while time-domain tests of exclusion may have limited power against dynamic mispecifications, frequency-domain tests are shown to perform rather satisfactorily. Higher-order asymptotic expansions of two

[4]Initial work on spectral methods for stationary time series includes, inter alia, Hannan (1963a, 1963b), Hannan (1970), Hannan and Thomson (1971), Hannan and Robinson (1973), Engle (1974), Engle (1978), and Harvey (1978).

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

alternative spectral regression estimators are developed in Xiao and Phillips (1998) and used to provide formulae for band(width) selection in spectrum regression.

Corbae, Ouliaris, and Phillips (2002) extend and refine the theory in Hannan (1963a, 1963b) by allowing for the presence of stochastic and deterministic trends. Among other results, Corbae, Ouliaris, and Phillips (2002) show that de-trending by removing deterministic components in the time domain, before applying band spectrum regression, is—in general—not equivalent to de-trending in the frequency domain and then running a band spectrum regression. Should there be variation in the slope coefficients across frequency bands, such variation ought to be taken into account for effective de-trending. Not doing so would lead to a sort of misspecification bias and potentially biased (and inconsistent, for nonstationary data) spectral estimates of the (frequency-specific) regression coefficients. Variation across frequencies, and its determinants in the time domain, will represent the core of our discussion. We, therefore, return to de-trending formally in Section 3.1.2.

Naturally, band spectrum regression has been used to estimate low-frequency, or long-run, relations between economic time series with stochastic trends, that is, co-integrating relations. In particular, Phillips (1991) provides a treatment in which Generalized Least-Squares (GLS) system spectrum regression, possibly restricted to a narrow band of frequencies around the origin, produces asymptotically efficient estimates of co-integrating vectors. In their rich discussion of inference for long-run equilibrium models with stochastic trends, Phillips and Loretan (1991) add to the work of Phillips (1991) by proposing an augmented spectral regression approach for estimating co-integrating vectors. The augmentation is intended to generate an asymptotically block-diagonal error structure in the co-integrating system and, therefore, an alternative way in which spectral estimates can be made as efficient as full-system maximum likelihood.[5] Related ideas are used by Choi and Phillips (1993) to construct nuisance parameter-free frequency-domain unit-root tests. More recently, Andersen and Varneskov (2021) develop a band spectrum approach for inference on the relation between (the weakly dependent errors of) ARFIMA processes in which each variable may have a different (fractional) integration order. Their procedure relies on (medium) band spectrum regression on series which, in the first step, have been purged of long memory.[6]

Multiresolution analysis has advocated methods of inference which localize in both time and scale, thereby facilitating interpretation over frequency-domain procedures (Gençay, Selçuk, and Whitcher, 2001, is a classical reference). There are interesting logical links, which we make explicit in Section 4, between methods of inference in the frequency domain and time-scale inferential methods. However, while the former have been formalized thoroughly, the latter have ubiquitously

[5]Both Phillips (1991) and Phillips and Loretan (1991) emphasize the usefulness of spectral methods over full-system maximum likelihood in light of their robustness to dynamic misspecification of the error terms.

[6]Earlier spectral methods for inference on types of fractional models were proposed, for example, in Marinucci (2000), Marinucci and Robinson (2001), Marinucci and Robinson (2003), Bandi and Perron (2006) and Christensen and Nielsen (2006).

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

been used as *filtering* procedures (Section 4.1). The nonparametric nature of these filters is appealing in applications which may dispense with attention to the structure of the underlying data generating process. It is less appealing in situations in which transitioning from time-scale representations to time-domain representations—a theme of this article—is statistically and economically revealing. A typical example is the use of *wavelets*, which we will discuss below at some length. Wavelets are generally employed to extract components with various levels of persistence from a time series of interest, in much the same way in which a Beveridge and Nelson (1981, hereafter BN) decomposition is designed to separate a parametric time-series model into a permanent and a transitory component. In light of their nonparametric nature, wavelets do so, however, without a formal mapping with familiar time-series models.[7]

A new time-scale representation of the classical Wold representation, dubbed *extended* Wold representation, is proposed and studied in the work of Bandi et al. (2019) and Ortu et al. (2020). This is a first step in the direction of bringing multiresolution methods into the fold of time-series econometrics. The result is a modeling approach which maps linear and nonlinear covariance-stationary time series into a suitable time-scale representation, thereby aiding inference (which may now rest upon a vast body of work in time series econometrics) and economic interpretation (through localization in both scale and time). We will return to the extended Wold representation in Sections 4.3 and 4.4.

2.1. Macroeconomics

Band spectrum estimation is ideally suited to evaluate economic relations which may not to be stable across frequencies. One early example is the marginal propensity to consume permanent income relative to transitory income. Returning to Friedman's work mentioned in Section 1 (Friedman, 1957), the idea that consumption should depend (linearly) on permanent income, and less so on transitory income, is tested and disputed in Engle (1974). Engle (1978) evaluates the elasticity of output prices with respect to wages. Such elasticity is found to be larger over low frequencies than over high frequencies. Corbae, Ouliaris, and Phillips (1994) apply the GLS-type system-wide spectral methods in Phillips (1991) to estimate the co-integrating vector between (permanent) consumption and (permanent) income. Because of endogeneity, the dependence between consumption and income away from the origin is, instead, estimated by virtue of a suitably defined band spectrum instrumental variable estimator. Their results reject unity as the value of the partial effect of income on consumption at zero frequency. They, however, support the existence of differential responses of consumption to

[7]The Hodrick–Prescott filter (Hodrick and Prescott, 1997) is a more classical nonparametric alternative to the BN decomposition. For a recent debate regarding its efficacy, and related methodological contributions, we refer the reader to, for example, Hamilton (2018), Hodrick (2020), Phillips and Jin (2020), and Phillips and Shi (2020).

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

transitory versus permanent income changes, that is, frequency-specific marginal propensities to consume.

In work related to band spectrum regression, Granger and Lin (1995) use the definition of frequency-specific interdependence provided by Hosoya (1991) for stationary processes in order to define notions of frequency-specific causality for processes with stochastic trends. Croux, Forni, and Reichlin (2001) propose a frequency-domain measure, named cohesion, of dynamic co-movement between time series. In the bi-variate case, the measure is naturally defined as the ratio between the co-spectrum and the square root of the product of the spectral densities. Restricting the measure to a set of frequency bands, and to the zero frequency in particular, is shown to provide implications for the study of synchronization in economic fluctuations and long-run co-integrating co-movements. In Croux, Forni, and Reichlin (2001), cohesion is employed to study the geography of business-cycle fluctuations within Europe and the United States.

The long-run co-movement between time series is at the heart of several macroeconomic theories. We have previously referred to the classical life-cycle model of consumption. Several other examples could be offered. The long-run purchasing power parity hypothesis, for instance, postulates that nominal exchange rates and relative nominal prices should move one-to-one over low frequencies. Similarly, long-run nominal interest rates should be related to long-run inflation rates through Fisher-type equations. These relations, and many others, provide motivation for the work of Müller and Watson (2018) on long-run co-variability. Their suggestion is to project the time series of interest onto suitably defined cosine functions. The resulting fitted values are smooth averages of the data capturing fluctuations over suitable low frequencies. Because of the orthogonality properties of the cosine function, the projection coefficients also have an interpretation in terms of cosine-weighted averages of the data. Importantly, the co-variation of the former (the fitted values) is one-to-one with the co-variation of the latter (the projection coefficients), an instance of helpful dimensionality reduction associated with a small number of data averages. In previous contributions, Müller and Watson (2008) and Müller and Watson (2016) focus on the scalar case with an emphasis on testing time-series models through the lens of their low-frequency behavior and characterizing low-frequency prediction accuracy, respectively. Müller and Watson (2017) review this line of work. Projections on trigonometric functions, as well more general orthonormal functions, have some history in econometrics in contexts other than the evaluation of low-frequency cycles: methods designed to model and approximate time series with orthonormal sequences are introduced and studied in Phillips (1998) and in Phillips (2005).

Finally, interesting ideas on the use of spectral techniques for evaluating possibly misspecified dynamic structural economic models are contained in the work of, for example, Watson (1993), Diebold, Ohanian, and Berkowitz (1998), and Christiano and Vigfusson (2003). In essence, these papers call for the assessment of goodness of fit in the frequency domain, that is, they call for examining how

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

accurately a parametric model fits the data spectrum across a specified range of frequencies.

2.2. Finance

The capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) assumes that all investors are single-period expected utility (of terminal wealth) maximizers. As is well-known, in the CAPM, systematic risk (i.e., beta) is a linear projection coefficient of excess asset returns on excess market returns. Importantly, the CAPM does not provide any particular restriction on the (common, across investors) length of time over which investors optimize utility. As a consequence, the model does not offer any meaningful empirical implication about whether rates of return, and betas, should be measured over a week, a month or any other length of time.

This is a nontrivial logical and applied issue. Working with raw (rather than logarithmic) returns assumed to be independent and identically distributed, Levhari and Levy (1977) show that the return covariances and variances do not vary, in general, proportionally with the time horizon. Hence, if there is an "equilibrium" horizon over which the implications of the CAPM should be tested, betas on returns over alternative horizons may be different from the "equilibrium betas" (see, also, Schwartz and Whitcomb, 1977; Smith, 1978). Handa, Kothari, and Wasley (1989) apply the logic in Levhari and Levy (1977) to the evaluation of the size effect. They conclude that the return horizon matters. In contrast to results obtained using (traditional) monthly returns, the size effect is attenuated by the use of annual returns in the computation of CAPM betas. One of the premises of Kothari, Shanken, and Sloan (1995) is that the classical findings of Fama and French (1992), which are famously unsupportive of the CAPM, may be due to the chosen (monthly) return frequency. In the same vein as Handa, Kothari, and Wasley (1989), but focusing on alternative portfolios, the results in Kothari, Shanken, and Sloan (1995) reveal economically and statistically significant compensations for market risk when using annual betas (see, also, Handa, Kothari, and Wasley, 1993). As the holding period increases, the variation in expected returns is driven by the variation in future price levels and asset pricing tests focusing on long-run returns effectively amount to tests on prices. Cohen, Polk, and Vuolteenaho (2009) make this point while providing evidence for the pricing ability of a low-frequency notion of the CAPM using long-run (buy-and-hold) returns (as well as long-run cash flows) to compute betas.

Several reasons have been put forward by prior research to explain the frequency-dependence of measured betas. We have talked about the work of Levhari and Levy (1977) who compute returns over various horizons by aggregating (i.e., multiplying) raw returns under the assumption that the highest frequency returns are independent and identically distributed. Using continuously-compounded returns instead, Hawawini (1983) shows that beta changes across horizons because of the existence of serial cross-correlations between the returns

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

of individual securities and those of the market (we will detail the approach of Hawawini, 1983, in Section 5). Serial cross-correlation, and therefore beta sensitivity to the estimation length, is routinely attributed to a variety of high-frequency frictions in the trading process (c.f. Cohen et al., 1983), asynchronous trading being the leading cause in the literature (e.g., Scholes and Williams, 1977, Dimson, 1979, and, more recently, Lo and MacKinlay, 1990). We will soon return to another driver of frequency-specific effects in beta estimation which was put forward in Section 1, that is, slow adjustments to information. Finally, there seems to exist a seasonal component in monthly returns (see, e.g., the work of Rozeff and Kinney, 1976 and Keim, 1983). As emphasized by Kothari, Shanken, and Sloan (1995), the use of low-frequency (e.g., annual) returns (in the computation of betas, for example) is one way of sidestepping the statistical complications that arise from this seasonality.

Implicit in the above discussion is that much of the finance work exploring dependence on frequency hinges on aggregation. Spectral analysis, however, does have some tradition (see, e.g., the work of Berkowitz, 2001 and Cogley, 2001). Sarno, Thornton, and Wen (2007) make use of a frequency-domain representation of a vector autoregression to study the frequency-specific effect of a Fed fund rate shock on U.S. interest rates with different maturities. Dew-Becker and Giglio (2016) develop a novel spectral representation for the prices of risk associated with shocks to priced variables. The representation is used to evaluate the importance placed by popular asset pricing models on alternative frequencies (interesting related work is contained in Otrok, Ravikumar, and Whiteman, 2002, and Yu, 2012). Chaudhuri and Lo (2018) employ spectral analysis to measure the skill of a portfolio manager across horizons. Their work links spectral analysis to performance evaluation, thereby enabling the study of the horizon properties of active management (see, also, Chaudhuri and Lo, 2015). Neuhierl and Varneskov (2021) use frequency-domain techniques to develop a permanent-transitory spectrum decomposition of the stochastic discount factor (SDF) in a rich setting that accommodates stochastic volatility in the SDF and non-Gaussian shocks. The proposed decomposition allows them to study the frequency-specific contributions of the transitory component of the SDF to the unconditional asset return premium.

Multiresolution analysis and its extensions are, arguably, less traditional. Gençay, Selçuk, and Whitcher (2003), Gençay, Selçuk, and Whitcher (2005) and, more recently, Kang, In, and Kim (2017) define wavelet-based scale-specific betas. Ortu, Tamoni, and Tebaldi (2013) make use of the (redundant) Haar wavelet (see, e.g. Shensa, 1992) to obtain time and scale-specific components of consumption growth and explain the market risk premium. Market return forecasts using wavelet filters are discussed in Faria and Verona (2018) and Faria and Verona (2020). Bandi et al. (2019) define a notion of "scale-wise predictability" based on a suitable time-scale decomposition of the traditional Wold representation, called earlier *extended* Wold representation, and show that classical predictability implies a restricted form of "scale-wise predictability." Their framework is designed to free up the implied restrictions of classical approaches to predictability. Using the extended

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Wold representation in Bandi et al. (2019) and Ortu et al. (2020), Bandi et al. (2021) show how to represent risk factors as sums of orthogonal components capturing fluctuations over different scales. The representation leads to novel *spectral factor models* in which systematic risk is allowed (without being forced) to vary across frequencies. Frequency-specific systematic risk is captured by a notion of *spectral beta*, one which is based on the structure of the Wold coefficients, and therefore on a data generating process, rather than on a wavelet filter as in, for example, Gençay, Selçuk, and Whitcher (2003). An empirical example of the methods in Bandi et al. (2021) was provided in Section 1. We will expand on these methods in what follows. Bandi and Tamoni (2021) study the cross-sectional pricing potential of a business-cycle component of consumption growth. Their work provides a theoretical treatment of the mapping between aggregation and time-scale methods (whether wavelet-based or based on the extended Wold representation), to which we will return in Section 5.

2.3. More on the Positioning of this Article: Frequency-Specific Effects in Finance

Consistent with our previous discussion, we will focus on a central object of econometric inference, namely a regression slope or, using finance jargon, a beta. Even though the finance literature has placed emphasis—particularly for individual stocks—on time-variation in betas,[8] we will dispense with time-variation and focus instead on frequency-specific effects, a dimension which is only now beginning to gain attention as a first-order effect both for individual stocks and for portfolios.

One may argue that the empirical finance literature developed between the early seventies and the mid-90s (cited above) was, in fact, devoted to this very issue and largely exhausted it. We, instead, emphasize that the early literature contributed immensely to our current understanding of the problem from the point of view of *measurement*. Not coincidentally, some of the intuition and findings at the heart of papers like those of Scholes and Williams (1977), Dimson (1979), and Lo and MacKinlay (1990), to recall just a few fundamental contributions, inform the inferential theory and computation of covariance measures (and related quantities, like betas) when using frictional (in particular, asynchronous) intra-daily data, that is, data not available at the time of those earlier contributions. Yet, what that early literature did not study was the frequency-specific nature of systematic risk viewed as a genuine economic phenomenon. More explicitly, the use of alternative return frequencies and aggregation schemes was intended as a way to better *measure* the unique beta of interest. The recent literature, instead, begins with the assumption that the beta of interest is not unique. Beta may change across frequencies, scales, or horizons—as in the BRK example in the Introduction—for reasons other than

[8]There is consensus that the factor loadings associated with portfolios are considerably more stable than those associated with individual stocks (see, e.g., Kelly, Pruitt, and Su, 2019 Lettau and Pelger, 2020).

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

trading frictions. Different frequency-specific betas may have different pricing implications and, in all likelihood, different economic drivers (c.f. Kamara et al., 2016; Bandi et al., 2021; Neuhierl and Varneskov, 2021).

Why focus on frequency-specific risk? Much of the current empirical finance literature is devoted to dimensionality reduction (e.g., Harvey, Liu, and Zhu, 2016, Kelly, Pruitt, and Su, 2019, Feng, Giglio, and Xiu, 2020, Freyberger, Neuhierl, and Weber, 2020, Kozak, Nagel, and Santosh, 2020, and Kim, Korajczyk, and Neuhierl, 2021). A conclusion has been reached that "order" ought to be brought to a plethora of factors (used in explaining risk compensations) and pricing models (defined by the number of factors and their nature). Against this backdrop, attention to frequency has the potential to be revealing about both hidden sources of risk and their economic determinants, thereby helping bring discipline. As an example, the inability of the market index to price cross-sectional returns is, arguably, not overly palatable given basic economic logic. Yet, from a pricing standpoint, a business-cycle component of the market beta (i.e., a spectral market beta capturing business-cycle fluctuations) is found to be considerably more effective than the aggregate market beta and, therefore, able to crowd-out less interpretable factors (Bandi et al., 2021). Econometrically, then, the extraction of a priced frequency-specific beta may be interpreted as signal extraction from a noisy beta measure, the latter being the subject of the literature on beta measurement discussed above. This signal extraction problem, cast in a regression framework, will be central to our discussion.

Finally, the drivers of lagged adjustments (leading to frequency-specific betas) should be viewed as being economically broader than the asynchronicities in the early work of Scholes and Williams (1977) and others. For instance, the role of slow information diffusion, as caused by, for example, opacity, neglect and limited attention, in determining alternative forms of delayed adjustments over short-term horizons is central to the work of Gilbert et al. (2014),[9] Hong and Stein (1999), Hou and Moskowitz (2005), Hong, Torous, and Valkanov (2007), and Cohen and Frazzini (2008), and goes beyond that of high-frequency frictions. Medium-term and long-term lagged adjustments—with likely different economic determinants—would, also, in general play a role. Heston and Sadka (2008), for instance, report on long-run seasonal autocorrelations in stock returns superimposed on short-term positive autocorrelations (momentum, c.f. Jegadeesh, 1990) and medium-term negative autocorrelations (reversals, c.f. Bondt, Werner, and Thaler, 1985). In this article, the relevant delayed adjustments—at all horizons—are those between a variable (an asset return, in the BRK example) and its assumed factor(s) (the market return, in the BRK example).

[9]In their framework, "opacity" (i.e., uncertainty about the effect of systematic news on firm value) is a separate phenomenon from trading frictions and behavioral biases. BRK is, in their view, an "opaque" firm whose returns respond to market news with a delay (as we have shown empirically in Panel (c) of Figure 1) and, as a consequence, whose high-frequency market beta underestimates the uncertainty risk induced by opacity.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

We conclude with the following observation. As emphasized, our focus is on modeling and operationalizing frequency dependence in unconditional betas. Considering time-variation and frequency-specific effects in beta estimation *jointly* is, however, an important avenue for future work, one which appears to be valuable, for example, for the purpose of evaluating conditional pricing models, particularly those in which the conditioning set is not in the immediate past (risk depends on aggregate shocks) and, therefore, frequency dynamics may play a meaningful statistical and economic role. The work of Bandi and Su (2022), in which conditional heteroskedasticity is added to vector autoregressions before implementing frequency-based decompositions, is a tractable, initial way to approach the problem. Much remains to be done.

We now turn to formalizations. In order to help the reader, we keep technical details to a minimum but attempt to provide a cohesive treatment nonetheless.

3. A MODELING FRAMEWORK: FREQUENCY DOMAIN

We largely follow the notation in Corbae, Ouliaris, and Phillips (2002). Let x_t ($t = 1, \ldots, T$) be a k-dimensional time-series vector generated by

$$x_t = \Pi_2^\top z_t + \tilde{x}_t, \tag{1}$$

where z_t is a $(p+1)$-dimensional deterministic sequence, $\tilde{x}_t$ is a zero-mean k-dimensional time-series vector and Π_2 is a $(p+1) \times k$ matrix of coefficients. The series x_t, therefore, has both a deterministic component (z_t) and a stochastic (latent) component ($\tilde{x}_t$). Regarding the deterministic component, we confine our discussion to time trend polynomials, the most common case in applied work, and set $z_t = (1, t, \ldots, t^p)$.

Similarly to x_t, the observed dependent variable y_t is assumed to have both time trends and a stochastic component. Specifically, y_t is linked to z_t and the latent $\tilde{y}_t$ by virtue of the specification

$$y_t = \Pi_1^\top z_t + \tilde{y}_t, \tag{2}$$

in which Π_1 is a $(p+1) \times 1$ matrix of coefficients. The component $\tilde{y}_t$ is related to $\tilde{x}_t$ and ε_t according to a (possibly two-sided) distributed lag specification, that is,

$$\tilde{y}_t = \sum_{k=-\infty}^{\infty} \beta_k^\top \tilde{x}_{t-k} + \varepsilon_t = \beta^\top(L)\tilde{x}_t + \varepsilon_t, \tag{3}$$

in which the transfer function of the filter $b(\omega) = \beta\left(e^{i\omega}\right) = \sum_{k=-\infty}^{\infty} \beta_k e^{i\omega k}$ is assumed to converge for all $\omega \in [-\pi, \pi]$. Equation (3), a lead-lag model in the time domain, will be central to our discussion.

The process $\tilde{x}_t$ is taken to be independent of the shock ε_t and, therefore, strictly exogenous. We follow Corbae, Ouliaris, and Phillips (2002) and allow for both stationary and non-stationary $\tilde{x}_t$.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Assumption 1 (stationary $\tilde{x}_t$)**.** $\zeta_t = \left(\varepsilon_t, \tilde{x}_t^\top\right)^\top$ is a jointly stationary time series with Wold representation $\zeta_t = \sum_{k=0}^{\infty} \widetilde{\boldsymbol{C}}_k \widetilde{\boldsymbol{\xi}}_{t-k} = \sum_{k=0}^{\infty} \begin{pmatrix} \widetilde{c}_k^1 & \widetilde{\boldsymbol{c}}_k^2 \\ \boldsymbol{c}_k^3 & \boldsymbol{c}_k^4 \end{pmatrix} \begin{pmatrix} \widetilde{\xi}_{t-k}^1 \\ \boldsymbol{\xi}_{t-k}^2 \end{pmatrix}$, with $\widetilde{\boldsymbol{c}}_k^2 = \boldsymbol{c}_k^3 = 0$ for all k, where the vector process $\widetilde{\boldsymbol{\xi}}_t$ is white noise mean zero and variance Σ_ξ with a finite fourth moment and the coefficients $\widetilde{\boldsymbol{C}}_k$ satisfy $\sum_{k=0}^{\infty} k^{\frac{1}{2}} \left\| \widetilde{\boldsymbol{C}}_k \right\| < \infty$.

In agreement with Assumption 1, the symbol ε will always be used to define exogenous regression errors (see, e.g., the example in Section 4.4). The tildes in the definition of ζ_t will be dropped when expressing the Wold representation of $\left(\tilde{y}_t, \tilde{x}_t^\top\right)^\top$ (see, again, Section 4.4) rather than that of $\left(\varepsilon_t, \tilde{x}_t^\top\right)^\top$.

Assumption 2 (nonstationary $\tilde{x}_t$)**.** $\tilde{x}_t$ is an I(1) process satisfying $\Delta \tilde{x}_t = w_t$, with $\tilde{x}_0$ given by any $O_p(1)$ random variable. The shocks $\zeta_t = \left(\varepsilon_t, w_t^\top\right)^\top$ satisfy Assumption 1.

It is now convenient to introduce the following notation. The discrete Fourier transform (DFT) of $\{a_t; t = 1, \ldots, T\}$ is $a_\omega = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} a_t e^{i\omega t}$, $\{\omega_j = \left(\frac{2\pi j}{T}\right) : j = 0, 1, \ldots, T-1\}$ are the fundamental frequencies, and $i = \sqrt{-1}$. Let $\tilde{X} = \left[\tilde{x}_1, \ldots, \tilde{x}_T\right]^\top$ be the $T \times k$ matrix of observations of the exogenous regressors in $\tilde{x}_t$. Also, let W be the $T \times T$ matrix with (j,t) element $w_{j,t} = e^{i\omega_j t}/\sqrt{T}$. The trigonometric identities ensure that the columns of W are orthonormal so that, denoting by W^* the complex conjugate transpose of the matrix W, we have $W^*W = I_T = WW^*$. We note that the DFT of $\tilde{X}$ is given by $\tilde{x}_{\boldsymbol{\omega}} = W\tilde{X} = [\tilde{x}_{\omega_0}, \ldots, \tilde{x}_{\omega_{T-1}}]^\top$. Let Z be the $T \times (p+1)$ observation matrix of the nonstochastic regressors z_t, let P_Z be the projection matrix onto the range of Z and let, also, $Q_Z = I_T - P_Z$ be the corresponding residual projection matrix. The same notation will be used for other variables and should, therefore, not cause confusion. For example, consistent with the definitions of $\tilde{X}$ and $\tilde{x}_{\boldsymbol{\omega}}$, we will write $\tilde{y}_{\boldsymbol{\omega}} = W\tilde{Y} = [\tilde{y}_{\omega_0}, \ldots, \tilde{y}_{\omega_{T-1}}]^\top$.

3.1. An Illustrative Case: Two Frequency-Specific Betas

As in Corbae, Ouliaris, and Phillips (2002), we provide an example in which the transfer function $b(\omega)$ is stable over two frequency bands. The two-beta example is a simple, yet revealing, case which will be employed repeatedly, and for alternative approaches, in this article.

Consider the bands $\mathcal{B}_A = [-\omega_0, \omega_0]$ and $\mathcal{B}_A^c = [-\pi, -\omega_0) \cup (\omega_0, \pi]$ for a given frequency $\omega_0 > 0$ and define the frequency-dependent partial effect

$$b(\omega) = \beta_A 1[\omega \in \mathcal{B}_A] + \beta_{A^c} 1\left[\omega \in \mathcal{B}_A^c\right], \tag{4}$$

where β_A is a k-vector of parameters corresponding to the frequency band $\mathcal{B}_A$ and β_{A^c} is the k-vector of parameters corresponding to the frequency band $\mathcal{B}_A^c$. Thus, β_A

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

captures partial effects (in a regression of $\tilde{y}$ onto $\tilde{x}$) over low frequencies ($|\omega| \leq \omega_0$) whereas β_{A^c} captures partial effects over high frequencies ($|\omega| > \omega_0$).

The following matrices will now be useful to characterize the band spectrum estimates. As in Engle (1974), define the $T \times T$ "selector" matrix A with diagonal elements equal to 1 corresponding to the frequencies in $\mathcal{B}_A$ and zeros elsewhere. Clearly, $A^c = (I_T - A)$ selects the remaining frequencies over $\mathcal{B}_A^c$. Both A and A^c are symmetric and idempotent (projection) matrices such that $A^c A = AA^c = 0$. Define, also, $\Psi = W^* A W$ and its analogue over $\mathcal{B}_A^c$, that is, $\Psi^c = W^* A^c W$. We note that $\Psi^c = I_T - \Psi$, by the unitary property of W.

The band spectrum regressions over the assumed frequency bands can now be written as

$$AW\tilde{y} = AW\tilde{X}\beta_A + AW\varepsilon, \tag{5}$$

$$A^c W\tilde{y} = A^c W\tilde{X}\beta_{A^c} + A^c W\varepsilon, \tag{6}$$

where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)^\top$.[10,11]

Recalling that $A + A^c = I_T$, adding equations (5) and (6), and premultiplying by the conjugate transpose W^* gives an alternative *time-domain specification* for $\tilde{y}$, that is,

$$\tilde{y} = \Psi\tilde{X}\beta_A + \Psi^c\tilde{X}\beta_{A^c} + \varepsilon, \tag{7}$$

$$= \tilde{X}_\Psi \beta_A + \tilde{X}_{\Psi^c}\beta_{A^c} + \varepsilon, \tag{8}$$

one which is separate from the original time-domain representation in equation (3), but will be of considerable interest for us. We note, in fact, that

$$\tilde{X} = \tilde{X}_\Psi + \tilde{X}_{\Psi^c}, \tag{9}$$

$$\tilde{X}_\Psi^\top \tilde{X}_{\Psi^c} = 0. \tag{10}$$

In other words, for any finite sample, the orthogonal components $\tilde{X}_\Psi$ and $\tilde{X}_{\Psi^c}$ capture, respectively, low-frequency and high-frequency dynamics in the exogenous variable $\tilde{x}$ while re-constructing it upon simple summation. This is a Fourier-based filtering procedure. We will return to it.

As Corbae, Ouliaris, and Phillips (2002) state: "In equation (7), $\left(\varepsilon_t, \tilde{x}_t^\top\right)^\top$ satisfies Assumption 1 in the stationary case and $\left(\varepsilon_t, \Delta\tilde{x}_t^\top\right)^\top$ satisfies Assumption 2 in the nonstationary case. Thus, equation (7) is a variable (band dependent) coefficient time series regression with exogenous regressors and stationary errors

[10] Early finite-sample tests of $H_0 : \beta_A = \beta_{A^c}$ are contained in Engle (1974) and Engle (1978).

[11] The exact mapping between equation (3) and equations (5) and (6), along with the order of the approximation term under both stationarity and non-stationarity, is discussed in Corbae, Ouliaris, and Phillips (2002), pp. 1071–1073. While in the stationary case the models are the same up to a vanishing term of probability order $\frac{1}{\sqrt{T}}$, this is not so in the nonstationary case.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

under Assumption 1 and a co-integrating regression with exogenous regressors and band dependent coefficients under Assumption 2."

Re-writing, now, equations (1) and (2), we have

$$X = Z\Pi_2 + \tilde{X},$$
$$y = Z\Pi_1 + \tilde{y},$$

respectively. Using equation (7), it follows that the model can be re-expressed as

$$y = \Psi Z(\Pi_1 - \Pi_2\beta_A) + \Psi^c Z\left(\Pi_1 - \Pi_2\beta_A^c\right) + \Psi X\beta_A + \Psi^c X\beta_{A^c} + \varepsilon. \quad \textbf{(11)}$$

Just like it is the case for the exogenous regressor $\tilde{X}$, the trend regressor Z now appears in the specification in equation (11) with frequency-dependent coefficients.

3.1.1. *Band Spectrum Regression Without Deterministic Trends.* We begin by considering the stationarity case (Assumption 1) and dispense with deterministic trends, that is, $Z = 1$. This is the model specification introduced by Hannan (1963a, 1963b) and popularized in economics by Engle (1974).

Dropping the tilde (because $\tilde{x}$ and $\tilde{y}$ are no longer latent) and re-writing equation (5), we obtain

$$Ay_{\boldsymbol{\omega}} = Ax_{\boldsymbol{\omega}}\beta_A + A\varepsilon_{\boldsymbol{\omega}}, \quad \textbf{(12)}$$

where we recall that $y_{\boldsymbol{\omega}} = Wy$, $x_{\boldsymbol{\omega}} = WX$, and $\varepsilon_{\boldsymbol{\omega}} = W\varepsilon$. Naturally, by the idempotency of A, the ordinary least squares (OLS) estimator of β_A is

$$\hat{\beta}_A = \left(x_{\boldsymbol{\omega}}^* A x_{\boldsymbol{\omega}}\right)^{-1} x_{\boldsymbol{\omega}}^* A y_{\boldsymbol{\omega}} = \left(X^\top \Psi X\right)^{-1} X^\top \Psi y. \quad \textbf{(13)}$$

An analogous expression applies to $\hat{\beta}_{A^c}$, namely, the slope estimator over the frequency band $\mathcal{B}_A^c$.

Remark 1. Under the assumption that X (and therefore $x_{\boldsymbol{\omega}}$) are independent of the error term, the Gauss-Markov theorem implies that the estimator in equation (13) is best linear unbiased. If the model applies to all frequencies and A is the identity matrix, this estimator is, of course, identical to the OLS estimator in $y = X\beta + \varepsilon$, a result which follows directly from the unitary property of W. In this case, if a band spectrum regression is run instead, the slope estimator remains unbiased but, as emphasized in Engle (1974), it will be relatively inefficient.

Remark 2 (c.f. Engle, 1974; Harvey, 1978)**.** Premultiplying equation (12) by W^* has the effect of "inverting" the band spectrum regression to the time domain:

$$W^*AWy = W^*AWX\beta_A + W^*AW\varepsilon. \quad \textbf{(14)}$$

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

The result is, in effect, a time-domain regression based on band-pass filters for the regressor and the regressand.[12] By using the structure of the matrix $\Psi = W^* A W$, it is easy to see that the time-domain estimate of β_A is identical to the frequency-domain estimate in equation (13).[13]

3.1.2. *Band Spectrum Regression with Deterministic Trends.* Hannan's recommendation for handling deterministic trends is to filter the data by a trend-removing regression via the projection matrix Q_Z (defined earlier) before running a band spectrum regression (c.f. Hannan, 1963a,b). In this case, the slope estimator over the frequency band $\mathcal{B}_A$ has the form

$$\hat{\beta}_A = \left(X^\top Q_Z \Psi Q_Z X\right)^{-1} X^\top Q_Z \Psi Q_Z y. \tag{15}$$

Again, an analogous expression applies to $\hat{\beta}_{A^c}$. Using equation (11), the fact that $\Psi^c = I_T - \Psi$, and equation (15), Corbae, Ouliaris, and Phillips (2002) show that

$$\mathbb{E}(\hat{\beta}_A | X) = \beta_A - \left(\tilde{X}^\top Q_Z \Psi Q_Z \tilde{X}\right)^{-1} \tilde{X}^\top Q_Z \Psi Q_Z \Psi^c \tilde{X} (\beta_A - \beta_{A^c}), \tag{16}$$

$$\mathbb{E}(\hat{\beta}_{A^c} | X) = \beta_{A^c} - \left(\tilde{X}^\top Q_Z \Psi Q_Z \tilde{X}\right)^{-1} \tilde{X}^\top Q_Z \Psi Q_Z \Psi \tilde{X} (\beta_{A^c} - \beta_A). \tag{17}$$

Hence, if $\beta_A \neq \beta_{A^c}$, band spectrum regression yields (conditionally) biased estimates of the slope coefficients when the trend-removing regression is conducted via the projection matrix Q_Z, as is customary in the time domain.[14] The case $\beta_A \neq \beta_{A^c}$ is, of course, of general interest and the results in Corbae, Ouliaris, and Phillips (2002) point to a misspecification bias when partial effects are frequency-specific. An analogous result arises in other contexts, as we will see in what follows.

As stressed by Corbae, Ouliaris, and Phillips (2002), the bias has a finite sample nature, and disappears with $T \to \infty$, for stationary $\tilde{x}_t$ (c.f. Assumption 1). Corbae, Ouliaris, and Phillips (2002), however, also show that when $\tilde{x}_t$ is integrated (as in Assumption 2), the bias may not vanish in the limit, thereby affecting inference

[12] Band-pass filters have a long tradition in applied macroeconomics, less so in empirical finance due to its more recent attention to dynamics over the business-cycle and lower frequencies. A recent use of a filter constructed as a two-sided moving average of suitable length is contained in Comin and Gertler (2006). Their focus is on cycles (dubbed "medium-term" cycles) between 2 and 200 quarters. As they point out, the traditional focus in the literature is, instead, on the use of band-pass filters to capture cycles between 2 and 32 quarters, that is, cycles representative of business-cycle, and higher, frequencies (Baxter and King, 1999; Christiano and Fitzgerald, 2003).

[13] Doing exact inference on equation (12) versus equation (14) may require care. If the number of included frequencies in the band spectrum regression in equation (12) is $T' < T$, OLS on the frequency-domain data will pick up the fact that only $T' - k$ degrees of freedom are left. However, in equation (14) there are still T time-domain observations and a standard package program will assume $T - k$ degrees of freedom. Thus, one needs to correct the standard errors for the lost degrees of freedom in equation (14). See Engle (1974) and Harvey (1978).

[14] Unbiased time-domain de-trending may be implemented by applying the projection matrix $Q_{\Psi Z}$ on $X_\Psi = \Psi X$ to define the estimator $\hat{\beta}_A = \left(X_\Psi^\top Q_{\Psi Z} X_\Psi\right)^{-1} X_\Psi^\top Q_{\Psi Z} y$. This procedure would amount to frequency-specific time-domain de-trending (through $Q_{\Psi Z}$) of a frequency-specific time-domain regressor $X_\Psi = \Psi X$.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

even with large sample sizes. They suggest trend removal in the frequency domain instead.[15]

3.2. Frequency-Specific Betas: A Time-Domain Justification

Which time-domain specification leads to frequency-specific betas in the present framework? The answer is a lead-lag structure between $\tilde{y}_t$ and $\tilde{x}_t$ consistent with equation (3). A simple way to verify this observation follows directly from the work of Corbae, Ouliaris, and Phillips (2002).

Consider the specification with two frequency bands in Section 3.1. The coefficients β_k in the frequency-specific partial effect $b(\omega) = \beta\left(e^{i\omega}\right) = \sum_{k=-\infty}^{\infty} \beta_k e^{i\omega k}$ of equation (4) are

$$\beta_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} b(\omega)e^{-i\omega k}d\omega = \begin{cases} \frac{\beta_A\omega_0}{\pi} + \frac{\beta_{A^c}}{\pi}(\pi-\omega_0), & \text{if } k=0, \\ \frac{\sin(k\omega_0)}{\pi k}(\beta_A - \beta_{A^c}), & \text{otherwise.} \end{cases} \tag{18}$$

The outcome of the above Fourier inversion is symmetrical lead-lag effects in the response of $\tilde{y}$ to $\tilde{x}$.[16] Naturally, these lead-lag effects generate substantial dynamic cross-correlations between $\tilde{y}_t$ and $\tilde{x}_t$. If $\beta_A = \beta_{A^c}$, all leads and lags are zero and $\beta_0 = \beta_A = \beta_{A^c}$ which is, of course, intuitive.

Equation (18) constitutes a preliminary formalization of an observation we made in Section 1. Variation across frequencies is, in general, the result of dynamic structures in the time domain. The frequency-dependence of the BRK spectral betas was, in fact, imputed to delayed adjustments of the BRK returns to market news.

In what follows, we will work with stationary time series (i.e., under Assumption 1 and absence of time trends). Consistently, we will use the notation y_t ad x_t in place of $\tilde{y}_t$ and $\tilde{x}_t$, respectively.

4. AN ALTERNATIVE MODELING FRAMEWORK: TIME-SCALE DOMAIN

This section discusses procedures to localize in both time and scale. Section 4.1 illustrates wavelet methods of signal extraction over different levels of resolution in the tradition of Mallat (1989). Section 4.2 is about a form of wavelet regression. In Sections 4.3–4.6, we turn to recent work whose purpose is to formalize an

[15] Applying equation (12) to y and X with deterministic trends, we have $AWy = AWZ(\Pi_1 - \Pi_2\beta_A) + AWX\beta_A + AW\varepsilon$. An unbiased estimator of β_A would, therefore, be $\beta_A = (X^\top W^* A Q_{AWZ} AWX)^{-1} X^\top W^* A Q_{AWZ} AWy$, c.f. Corbae, Ouliaris, and Phillips (2002), equation (44).

[16] The case $k=0$ is immediately seen: $\beta_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} b(\omega)d\omega = \frac{\beta_A}{2\pi}\int_{-\omega_0}^{\omega_0} d\omega + \frac{\beta_{A^c}}{2\pi}\int_{\omega_0}^{\pi} d\omega + \frac{\beta_{A^c}}{2\pi}\int_{-\pi}^{-\omega_0} d\omega = \frac{\beta_A\omega_0}{\pi} + \frac{\beta_{A^c}(\pi-\omega_0)}{\pi}$. The case $k \neq 0$ just makes use of a suitable version of Euler's formula: $\beta_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} b(\omega)e^{-i\omega k}d\omega = \frac{\beta_A}{2\pi}\int_{-\omega_0}^{\omega_0} e^{-i\omega k}d\omega + \frac{\beta_{A^c}}{2\pi}\int_{\omega_0}^{\pi} e^{-i\omega k}d\omega + \frac{\beta_{A^c}}{2\pi}\int_{-\pi}^{-\omega_0} e^{-i\omega k}d\omega = \frac{\beta_A}{2\pi}\left|\frac{-e^{-i\omega k}}{ik}\right|_{-\omega_0}^{\omega_0} + \frac{\beta_{A^c}}{2\pi}\left|\frac{-e^{-i\omega k}}{ik}\right|_{\omega_0}^{\pi} + \frac{\beta_{A^c}}{2\pi}\left|\frac{-e^{-i\omega k}}{ik}\right|_{-\pi}^{-\omega_0} = \frac{\beta_A}{2\pi}\left(\frac{e^{i\omega_0 k}-e^{-i\omega_0 k}}{ik}\right) + \frac{\beta_{A^c}}{2\pi}\left(\frac{e^{-i\omega_0 k}-e^{-i\pi k}}{ik}\right) + \frac{\beta_{A^c}}{2\pi}\left(\frac{e^{i\pi k}-e^{i\omega_0 k}}{ik}\right) = \frac{\sin(k\omega_0)}{\pi k}(\beta_A - \beta_{A^c})$ given $e^{i\omega k} - e^{-i\omega k} = 2i\sin(k\omega)$ and $\sin(k\pi) = 0$ for all k.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

operational map between time-series modeling for covariance stationary processes and multiresolution methods (Bandi et al., 2019; Ortu et al., 2020). As in the previous section, our focus will be on linear dependence (Bandi et al., 2021).

4.1. Multiresolution Filters: Haar

Let x_t be a stationary time series of interest. By adding and subtracting the quantity $\frac{x_{t-1}}{2}$, we may write

$$x_t = \underbrace{\frac{x_t - x_{t-1}}{2}}_{\widehat{x}_t^{(1)}} + \underbrace{\left[\frac{x_t + x_{t-1}}{2}\right]}_{\pi_t^{(1)}}, \tag{19}$$

which effectively amounts to separating the series into a (sort of) "transitory" component $\widehat{x}_t^{(1)}$ and a (sort of) "persistent" component $\pi_t^{(1)}$. This is a Haar wavelet decomposition of the original process (Haar, 1910). Let us now add and subtract $\frac{x_{t-2}+x_{t-3}}{4}$ to break the "persistent" component in equation (19) into additional "persistent" and "transitory" components:

$$x_t = \underbrace{\frac{x_t - x_{t-1}}{2}}_{\widehat{x}_t^{(1)}} + \underbrace{\left[\frac{x_t + x_{t-1} - x_{t-2} - x_{t-3}}{4}\right]}_{\widehat{x}_t^{(2)}} + \underbrace{\left[\frac{x_t + x_{t-1} + x_{t-2} + x_{t-3}}{4}\right]}_{\pi_t^{(2)}}. \tag{20}$$

The procedure can, of course, be iterated further and generalized. To this extent, define the moving averages $\pi_t^{(j)}$ of length 2^j, namely,

$$\pi_t^{(j)} = \frac{1}{2^j}\sum_{k=0}^{2^j-1} x_{t-k}. \tag{21}$$

Clearly, $\pi_t^{(0)} = x_t$. These moving averages satisfy the iteration

$$\pi_t^{(j)} = \frac{\pi_t^{(j-1)} + \pi_{t-2^{j-1}}^{(j-1)}}{2}. \tag{22}$$

Next, express $\widehat{x}_t^{(j)}$ as the difference between two moving averages of length 2^{j-1} and 2^j, that is,

$$\widehat{x}_t^{(j)} = \pi_t^{(j-1)} - \pi_t^{(j)}. \tag{23}$$

We refer to $\widehat{x}_t^{(j)}$, with $j = 1, \ldots, J$, as to the jth component of the original time series x_t. Intuitively, $\widehat{x}_t^{(j)}$ captures fluctuations between 2^{j-1} and 2^j periods. It is readily observed that $\widehat{x}_t^{(1)}$, $\widehat{x}_t^{(2)}$, and $\pi_t^{(2)}$ coincide, for example, with the expressions in equation (20).

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Since $\pi_t^{(0)} = x_t$, computing the telescopic sum up over j of the terms in equation (23), and adding $\pi_t^{(J)}$, reconstructs the process:

$$x_t = \sum_{j=1}^{J} \widehat{x}_t^{(j)} + \pi_t^{(J)}, \tag{24}$$

for any $J \geq 1$. Equation (24) decomposes the time series x_t into the sum of components with half-life belonging to a specific (dyadic) interval in addition to a residual quantity $\pi_t^{(J)}$ representing a long-run average. The larger J, the closer the residual quantity is to a constant term. Importantly, the components may be employed to evaluate dynamics (as well as dependencies) over different frequencies. Several observations are in order.

Remark 3. The same decomposition is easily conducted by virtue of an orthonormal linear operator discussed in the work of Mallat (1989).

Remark 4. The decomposition does not lead to exactly orthogonal components, even though the "de-correlation" property of wavelets is well-known, even for highly persistent processes (c.f. Chapter 5 of Gençay, Selçuk, and Whitcher, 2001). In this sense, the analogous Fourier decomposition in equation (9) has superior finite-sample properties. By construction, in fact, the two components ($\tilde{X}_{\Psi}$ and $\tilde{X}_{\Psi^c}$) are guaranteed to be orthogonal in any finite sample.

Remark 5. A Haar filter is a convenient choice, but alternative filters could have been used instead (see, e.g., Daubechies, 1992). As discussed in Bandi and Tamoni (2021), the Haar filter has, at least, two properties which make it appealing in macroeconomics and macrofinance. First, because the typical length of the business cycle is 2–8 years, the dyadic nature of its structure (with components capturing fluctuations between 2^{j-1} and 2^j periods) is particularly useful to study business-cycle dynamics and, as a consequence, dynamics over higher, or lower, frequencies than the business cycle by virtue of quarterly data. Second, since the Haar filter is an aggregation filter, it is convenient to formalize the mapping between *frequency* and *aggregation*. Section 5 expands on this observation.

4.2. The Haar Filter: Spectral Regression

We now turn to dependence. We illustrate, again, by way of a two-beta representation (as in Section 3.1) and, for simplicity, using a Haar filter with $J = 1$, as in equation (19).

Ignoring error terms, consider a process y_t linked to a "high-frequency" component and a "low-frequency" component of x_t:

$$y_t = \beta_{HF} \widehat{x}_t^{(1)} + \beta_{LF} \pi_t^{(1)}. \tag{25}$$

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Define, now, $\Lambda_{LF} = \beta_{LF} - \beta_{HF} \gtreqless 0$. It is immediate to observe that y_t can also be written in terms of lags of x_t. Specifically, we have

$$\begin{aligned} y_t &= \beta_{HF}\widehat{x}_t^{(1)} + (\beta_{HF} + \Lambda_{LF})\,\pi_t^{(1)} \\ &= \beta_{HF}\left(\widehat{x}_t^{(1)} + \pi_t^{(1)}\right) + \Lambda_{LF}\pi_t^{(1)} \\ &= \beta_{HF}x_t + \Lambda_{LF}\left[\frac{x_t + x_{t-1}}{2}\right] \qquad (26) \\ &= \left(\beta_{HF} + \frac{\Lambda_{LF}}{2}\right)x_t + \frac{\Lambda_{LF}}{2}x_{t-1}. \qquad (27) \end{aligned}$$

Remark 6. Equation (26) is reminiscent of the heterogeneous autoregressive (HAR) model of Corsi (2009), i.e.,

$$RV_t = a + \beta_d RV_{t-1} + \beta_w RV_{t-5:t-1} + \beta_m RV_{t-22:t-1} + \varepsilon_t,$$

a model in which past daily (RV_{t-1}), weekly ($RV_{t-5:t-1}$), and monthly ($RV_{t-22:t-1}$) realized variance measures are used to forecast future daily realized variance.

Remark 7. Equation (27) is a distributed lag model in which one delayed adjustment affects the current value of the regressand. Consistent, again, with the illustration in Section 1, heterogeneous dependence over frequencies implies lagged dynamics in the predictor.

4.3. Extended Wold Representation: The Scalar Case

We now employ the structure of the Haar filter as a way to model time-scale dynamics for covariance-stationary time series. Specifically, we transition from the *nonparametric* (signal-extraction) nature of multiresolution analysis, as discussed in the previous subsection, to a *parametric*, model-based, approach in line with methods of time-series modeling and inference. The discussion follows the work of Bandi et al. (2019) and Ortu et al. (2020). To this work, we add details on interpretation and new perspectives.

Let x_t be a (mean zero, without loss of generality) stationary scalar process defined on the usual probability space $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Its Wold representation is such that there exists a zero-mean white noise process ξ_t so that, for any t,

$$x_t = \sum_{k=0}^{\infty} c_k \xi_{t-k}, \qquad (28)$$

where the equality is in the $\mathbb{L}^2$-norm and $\{c_k\}_{k\in\mathbb{N}_0}$ is a square-summable sequence of real (projection) coefficients. As compared to Assumption 1, we have simply dropped the superscript from shocks and coefficients to streamline notation.

In agreement with equation (24), the following representation of x_t holds:

$$x_t = \sum_{j=1}^{J} x_t^{(j)} + \pi_t^{(J)}, \qquad (29)$$

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

where

$$x_t^{(j)} = \sum_{k=0}^{\infty} \psi_k^{(j)} \xi_{t-k2^j}^{(j)}, \quad \pi_t^{(J)} = \sum_{k=0}^{\infty} \gamma_k^{(J)} \bar{\xi}_{t-k2^J}^{(J)}, \tag{30}$$

in $\mathbb{L}^2$, with

$$\psi_k^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} c_{k2^j+i} - \sum_{i=0}^{2^{j-1}-1} c_{k2^j+2^{j-1}+i} \right), \tag{31}$$

$$\xi_t^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \xi_{t-i} - \sum_{i=0}^{2^{j-1}-1} \xi_{t-2^{j-1}-i} \right), \tag{32}$$

and

$$\bar{\xi}_t^{(J)} = \frac{1}{\sqrt{2^J}} \sum_{i=0}^{2^J-1} \xi_{t-i}, \qquad \gamma_k^{(J)} = \frac{1}{\sqrt{2^J}} \sum_{i=0}^{2^J-1} c_{k2^J+i}.$$

The individual components $x_t^{(j)}$ have Wold representations defined with respect to the shocks $\xi_t^{(j)}$. These shocks are discrete Haar transforms (DHT) of the original shocks (equation (32)) and are white noise on the dilated support $S_t^{(j)} = \{t - k2^j : k \in \mathbb{N}_0\}$. Similarly, the parameters of the component-specific Wold representations of each $x_t^{(j)}$ are DHTs of the original parameters (equation (31)). They can naturally be viewed as the coefficients resulting from an $\mathbb{L}^2$-projection of x_t onto the space spanned by the scale-wise shocks $\xi_t^{(j)}$.

Importantly, given the structure of the shocks, it is simple to verify that the components are orthogonal, for all t. Hence, differently from equation (24), equation (29) is an exact orthogonal decomposition into scale-specific components (capturing dyadic cycles) and a long-run term $\pi_t^{(J)}$. As $J \to \infty$, the long-run term vanishes, thereby giving

$$x_t = \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} \psi_k^{(j)} \xi_{t-k2^j}^{(j)},$$

in $\mathbb{L}^2$, that is, an *extended* Wold representation.

An example: the AR(1) case. Assume x_t is a mean-zero autoregressive process of order 1 with autocorrelation $|\rho| < 1$ and shocks ξ_t. Using the lag operator L in the usual way, write

$$x_t = (1 - \rho L)^{-1} \xi_t = \xi_t + \sum_{k=1}^{\infty} \rho^k \xi_{t-k} = \sum_{k=0}^{\infty} c_k \xi_{t-k},$$

where $c_k = \rho^k$, giving the process' Wold representation. We now break x_t into orthogonal components. Fix a generic scale level $j \in \mathbb{N}$. The expressions of the

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

coefficients $\psi_k^{(j)}$ can be obtained by virtue of equation (31):

$$\psi_k^{(j)} = \frac{\rho^{k2^j}\left(1-\rho^{2^{j-1}}\right)^2}{\sqrt{2^j}(1-\rho)}$$

for any $k \in \mathbb{N}_0$. The component processes $x_t^{(j)}$ are, therefore,

$$x_t^{(j)} = \frac{\left(1-\rho^{2^{j-1}}\right)^2}{\sqrt{2^j}(1-\rho)} \sum_{k=0}^{\infty} \rho^{k2^j} \xi_{t-k2^j}^{(j)}.$$

Barring the proportionality term $\frac{\left(1-\rho^{2^{j-1}}\right)^2}{\sqrt{2^j}(1-\rho)}$, the structure of the coefficients $\psi_k^{(j)}$ is that of an AR(1) process defined on the support $S_t^{(j)} = \{t - k2^j : \quad k \in \mathbb{N}_0\}$. In essence, we may rewrite the original AR(1) as a (weighted) infinite sum of AR(1)s with time steps 2^j and autoregressive coefficients given by ρ^{2^j}, i.e.,

$$x_t = \sum_{j=1}^{\infty} \frac{\left(1-\rho^{2^{j-1}}\right)^2}{\sqrt{2^j}(1-\rho)} \times x_t^{(j)},$$

where

$$x_t^{(j)} = \rho^{2^j} x_{t-2^j}^{(j)} + \xi_t^{(j)}.$$

Remark 8. The use of Haar wavelets and, more generally, multiresolution analysis has provided a set of valuable tools to extract signal about dynamics over different horizons. The recent convergence between multiresolution analysis and time-series analysis (a convergence which relies—as discussed above—on the application of wavelets to time-series shocks rather than to data) is now offering novel possibilities in terms of modeling. The traditional approach is nonparametric in nature. The new approach is instead parametric and hinges on the identification of Wold shocks (and Wold coefficients) for the evaluation of the scale-wise components. In this sense, they could serve different but, in many cases, complementary purposes. The latter allows for more structural explorations that only a "model" would permit. The former operates as a nonparametric filter. Should the data-generating process be assumed to be the extended Wold, a Haar filter would not reproduce the Wold time-scale components ($x_t^{(j)}$) unless the process is white noise. However, it would offer a set of easy-to-use and informative *descriptive* tools which could assist parametric modeling, when parametric modeling is needed, or allow the researcher to dispense with a formal data-generating process, when a less structured approach is desirable (c.f. Bandi and Tamoni, 2021).

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

4.4. Extended Wold Representation: The Multivariate Case

Let $\mathbf{x}_t = (y_t, x_t^\top)^\top$ be a covariance-stationary vector process of dimension $k+1$. Without loss of generality, again, we assume the process is mean zero.

Define the white noise process $\boldsymbol{\xi}_t = (\xi_t^1, \boldsymbol{\xi}_t^{2\top})^\top$ such that $\mathbb{E}[\boldsymbol{\xi}_t] = 0$ and $\mathbb{E}[\boldsymbol{\xi}_t \boldsymbol{\xi}_t^\top] = \Sigma_\xi$, where Σ_ξ is of dimension $(k+1, k+1)$. For any t, $(y_t, x_t^\top)^\top$ satisfies the following Wold representation:

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \sum_{k=0}^{\infty} \begin{pmatrix} c_k^1 & \boldsymbol{c}_k^2 \\ \boldsymbol{c}_k^3 & \boldsymbol{c}_k^4 \end{pmatrix} \begin{pmatrix} \xi_{t-k}^1 \\ \boldsymbol{\xi}_{t-k}^2 \end{pmatrix} = \sum_{k=0}^{\infty} \boldsymbol{C}_k \boldsymbol{\xi}_{t-k}, \tag{33}$$

with $\boldsymbol{C}_0 = I_{(k+1)}$, where the equality is in the $\mathbb{L}^2$-norm.

Suitable aggregation of the system shocks as in equation (32) leads to a representation in which the components $\mathbf{x}^{(j)}$ are orthogonal, for all t:

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} \boldsymbol{\Psi}_k^{(j)} \boldsymbol{\xi}_{t-k2^j}^{(j)} = \sum_{j=1}^{\infty} \mathbf{x}_t^{(j)}. \tag{34}$$

In this representation, the shocks $\boldsymbol{\xi}_t^{(j)}$, i.e.,

$$\boldsymbol{\xi}_t^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \boldsymbol{\xi}_{t-i} - \sum_{i=0}^{2^{j-1}-1} \boldsymbol{\xi}_{t-2^{j-1}-i} \right), \tag{35}$$

and the coefficient matrices $\boldsymbol{\Psi}_k^{(j)}$, i.e.,

$$\boldsymbol{\Psi}_k^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \boldsymbol{C}_{k2^j+i} - \sum_{i=0}^{2^{j-1}-1} \boldsymbol{C}_{k2^j+2^{j-1}+i} \right), \tag{36}$$

are, as earlier in the scalar case, the DHTs of the original Wold shocks and coefficients, respectively.

Bandi et al. (2021) show that the slope coefficient in a regression of y_t on an individual element of the k-vector x_t, say x_{ht}, can be expressed as follows:

$$\beta_h = \sum_{j=1}^{\infty} v_h^{(j)} \beta_h^{(j)}, \tag{37}$$

where $v_h^{(j)}$ is an "information ratio" given by the variance of the jth component of the regressor x_{ht} relative to the overall variance of the regressor and $\beta_h^{(j)} = \frac{\mathbb{C}(y_t^{(j)}, x_{ht}^{(j)})}{\mathbb{V}(x_{ht}^{(j)})}$ is a *spectral beta* corresponding to the jth component. In other words, it is a general result that the overall beta is a weighted average of spectral betas. The

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

BRK example in the Introduction provides estimates for seven spectral betas as well as for the corresponding overall beta.[17]

Equation (37) can easily be justified in a regression context. For simplicity and without loss of generality, let us assume that the regressor is the first element (i.e., $x_{ht} = x_{1t}$) of the k-vector x_t. Let us, also, assume that x_{1t} is broken down into only two components, namely the first one (i.e., $x_{1t}^{(1)}$) and a residual (orthogonal) component capturing lower frequency cycles (i.e., $x_{1t}^{(>1)}$). Clearly, $x_{1t}^{(>1)} = x_{1t} - x_{1t}^{(1)} = \sum_{j=2}^{J} x_t^{(j)} + \pi_t^{(J)}$, given equation (29).

Now, write

$$y_t = \beta_1^{(1)} x_{1t}^{(1)} + \beta_1^{(>1)} x_{1t}^{(>1)} + \varepsilon_t, \tag{38}$$

where ε_t is a shock uncorrelated with the components. Similarly to x_t, y_t can also be decomposed into two components, that is, $y_t = y_t^{(1)} + y_t^{(>1)}$. Because of the orthogonality of the components of y_t and x_{1t} *within* process, which is an implication of the Wold representation in equation (34), we have

$$\beta_1^{(1)} = \frac{\mathbb{C}(y_t, x_{1t}^{(1)})}{\mathbb{V}(x_{1t}^{(1)})}, \quad \beta_1^{(>1)} = \frac{\mathbb{C}(y_t, x_{1t}^{(>1)})}{\mathbb{V}(x_{1t}^{(>1)})}. \tag{39}$$

Because of the orthogonality of the components of y_t and x_{1t} *across* processes, which is also an implication of the Wold representation in equation (34), we have

$$\beta_1^{(1)} = \frac{\mathbb{C}(y_t, x_{1t}^{(1)})}{\mathbb{V}(x_{1t}^{(1)})} = \frac{\mathbb{C}(y_t^{(1)}, x_{1t}^{(1)})}{\mathbb{V}(x_{1t}^{(1)})}, \quad \beta_1^{(>1)} = \frac{\mathbb{C}(y_t, x_{1t}^{(>1)})}{\mathbb{V}(x_{1t}^{(>1)})} = \frac{\mathbb{C}(y_t^{(>1)}, x_{1t}^{(>1)})}{\mathbb{V}(x_{1t}^{(>1)})}, \tag{40}$$

but these are precisely the definitions of the spectral betas $\beta^{(1)}$ and $\beta^{(>1)}$. Now, given equation (38), we obtain

$$\begin{aligned}
\beta_1 &= \frac{\mathbb{C}(y_t, x_{1t})}{\mathbb{V}(x_{1t})} \\
&= \frac{\mathbb{C}(\beta_1^{(1)} x_{1t}^{(1)} + \beta_1^{(>1)} x_{1t}^{(>1)} + \varepsilon_t, x_{1t}^{(1)} + x_{1t}^{(>1)})}{\mathbb{V}(x_{1t})} \\
&= v_1^{(1)} \beta_1^{(1)} + v_1^{(>1)} \beta_1^{(>1)},
\end{aligned}$$

where $v_1^{(1)} = \frac{\mathbb{V}(x_{1t}^{(1)})}{\mathbb{V}(x_{1t})}$ and $v_1^{(>1)}$ is defined analogously, which is a two-component version of the expression in equation (37).

Remark 9. We note that equation (38), and its extensions with more than two components, can be viewed as a generalized regression model in which the partial effect of x_t on y_t changes across components. In this sense, the classical univariate regression model is a restricted version of equation (38), one in which the partial

[17] Given the Wold representation and its extended counterpart, both the overall beta and the spectral betas can—of course—be calculated as a function of the Wold coefficients (Bandi et al., 2021, Corollary to Theorem 1). Remarks 11 and 16 will provide closed-form expressions for the spectral betas of two revealing models.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

effects are stable across components: $\beta_1^{(1)} = \beta_1^{(>1)} = \beta_1$ in our example. Models that are genuinely multivariate in both the number of regressors and the number of components can now be readily defined.

Remark 10. Equation (38) is akin to the *alternative time-domain specification* in equation (7). Similarly to equation (7), the regressors are orthogonal and, in spite of the time-domain formulation, we can interpret the slope estimates directly in terms of frequency-specific betas. Contrary to equation (7), however, we may view equation (38) as a data-generating process, the orthogonality of the components being the result of the Haar transform applied to the original Wold shocks rather than the result of the Fourier transform applied to the data. The orthogonality of the components can be revealing both in, for example, applied macroeconomics (a field which has historically relied on alternative methods, like band-pass filters) and finance. Comin and Gertler (2006), for instance, have emphasized that their focus on fluctuations between 2 and 200 quarters is partly motivated by the difficulty in orthogonalizing cycles between 2 and 32 quarters and cycles between 32 and 200 quarters (Comin and Gertler, 2006, p. 526). The usefulness of orthogonality in testing asset pricing models in which the factors have cycles of alternative lengths is discussed in Bandi et al. (2021).

An example: regression with AR(1) regressors. Write

$$y_t = \beta x_t + u_t, \tag{41}$$

with

$$x_t = \rho x_{t-1} + \xi_t, \tag{42}$$

a mean-zero AR(1) process, as in the previous example. The corresponding Wold representation is

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \sum_{k=0}^{\infty} \boldsymbol{C}_k \boldsymbol{\xi}_{t-k} = \sum_{k=0}^{\infty} \begin{pmatrix} 0 & \beta\rho \\ 0 & \rho \end{pmatrix}^k \begin{pmatrix} \xi_{t-k}^1 \\ \xi_{t-k}^2 \end{pmatrix}, \tag{43}$$

where

$$\begin{pmatrix} \xi_t^1 \\ \xi_t^2 \end{pmatrix} = \begin{pmatrix} u_t + \beta\xi_t \\ \xi_t \end{pmatrix},$$

and

$$\Sigma_\xi = \begin{pmatrix} \sigma_u^2 + \beta^2\sigma_\xi^2 + 2\beta\sigma_{u,\xi} & \sigma_{u,\xi} + \beta\sigma_\xi^2 \\ \bullet & \sigma_\xi^2 \end{pmatrix}$$

is the variance matrix of the system's shocks.

If $\sigma_{u,\xi} = \mathbb{C}(u_t, \xi_t) = 0$, there is no endogeneity and the spectral betas satisfy $\beta^{(j)} = \beta = \frac{\mathbb{C}(y_t, x_t)}{\mathbb{V}(x_t)}$ for all j. Assume, now, $\sigma_{u,\xi} \neq 0$. In this case, the regressor is endogenous and $\beta \neq \frac{\mathbb{C}(y_t, x_t)}{\mathbb{V}(x_t)}$. In fact, it is also the case that the spectral betas change across frequencies. Below, we provide a visual illustration. Before doing so, it is

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

insightful to write the linear projection

$$u_t = \frac{\sigma_{u,\xi}}{\sigma_\xi^2}\xi_t + \varepsilon_t, \tag{44}$$

where ε_t is a white noise error term uncorrelated with ξ_t. Plugging equation (44) into equation (41), we obtain

$$y_t = \left(\beta + \frac{\sigma_{u,\xi}}{\sigma_\xi^2}\right)x_t - \frac{\sigma_{u,\xi}}{\sigma_\xi^2}\rho x_{t-1} + \varepsilon_t. \tag{45}$$

If $\sigma_{u,\xi} \neq 0$, equation (45) is a delayed adjustment model in which y_t responds to present (x_t) and past (x_{t-1}) values of the regressors. Once more, there is an important, albeit sometimes not completely salient, correspondence between changing dependence over frequencies and delayed adjustment(s) in the corresponding time-series representation. In this case, lack of salience is induced by the endogeneity of an autoregressive regressor.

Next, we plot the spectral betas (as well as the spectral covariances and the spectral variances of the regressor) for the model in equations (41) and (42) using the following parameters: $\sigma_\xi = \sigma_u = 1, \sigma_{u,\xi} = -0.3$ (varied to $\sigma_{u,\xi} = 0$) and $\rho = 0.5$ (varied to $\rho = 0.1$ and $\rho = 0.8$). Given the system's Wold representation in equation (43), the spectral betas, the spectral covariances and the spectral variances of the regressor are computed by using the Wold-based expressions in the Corollary to Theorem 1 of Bandi et al. (2021). Remark 11 formalizes the spectral betas in closed form and Figure 3 provides a visualization.

Remark 11. Using the Corollary to Theorem 1 of Bandi et al. (2021), it is easy to express the spectral betas for the model specification in equations (41) and (42) in closed form. Define

$$a_j = \sum_{i=1}^{2^{(j-1)}} \rho^{(i-1)},$$

$$b_j = a_j(1-\rho^{2^{(j-1)}}),$$

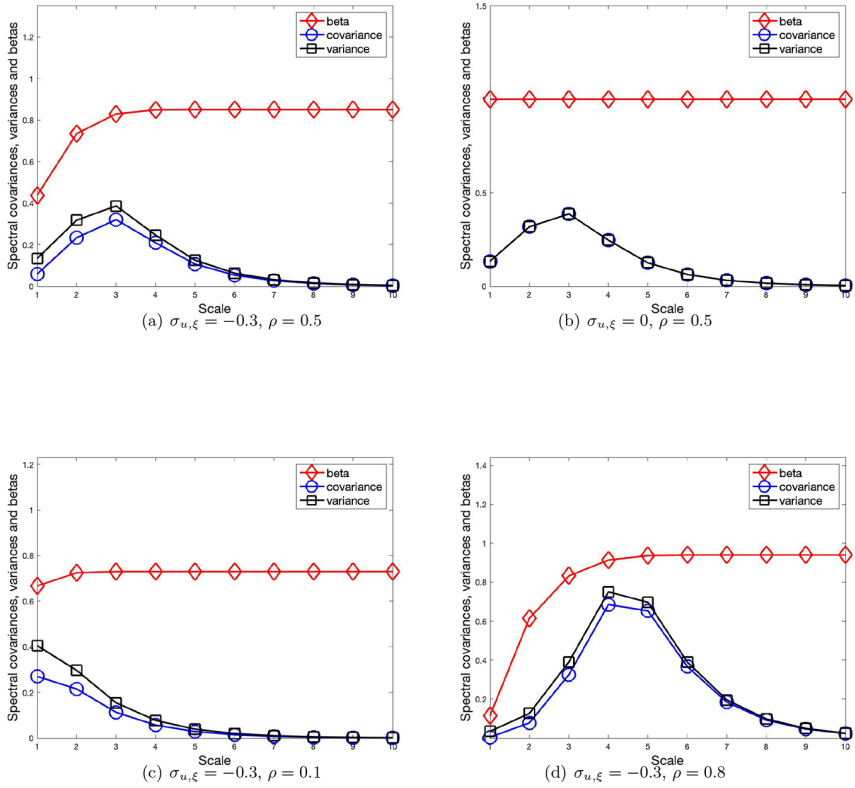
$$c_j = \sum_{i=0}^{2^{(j-1)}-1} \rho^i,$$

$$d_j = \rho(c_j - \rho^{2^{(j-1)}-1} - (\rho^{2^{(j-1)}-1})c_j).$$

We have:

$$\beta^{(j)} = \frac{(\sigma_{u,\xi} + \beta\sigma_\xi^2)b_j + \sigma_\xi^2\beta d_j b_j + \sigma_\xi^2\beta b_j^2(\frac{1}{1-\rho^{4j}} - 1)}{\sigma_\xi^2 b_j^2(\frac{1}{1-\rho^{4j}})}.$$

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press



(a) $\sigma_{u,\xi} = -0.3$, $\rho = 0.5$ (b) $\sigma_{u,\xi} = 0$, $\rho = 0.5$

(c) $\sigma_{u,\xi} = -0.3$, $\rho = 0.1$ (d) $\sigma_{u,\xi} = -0.3$, $\rho = 0.8$

FIGURE 3. The spectral betas in a regression model. We represent spectral variances, spectral covariances and spectral betas for the model in equations (41) and (42). We consider two values of the correlation between the error terms ($\sigma_{u,\xi} = -0.3$ in Panels (a), (c), (d), and $\sigma_{u,\xi} = 0$ in Panel (b)) and three values for the persistence of the factor ($\rho = 0.5$ in Panels (a), (b), $\rho = 0.1$ in Panel (c) and $\rho = 0.8$ in Panel (d)).

We observe that, when $\rho > 0$, a negative covariance $\sigma_{u,\xi}$ gives rise to positive delayed adjustments (because $\frac{\sigma_{u,\xi}}{\sigma_\xi^2}\rho$ in equation (45) is positive) and increasing, albeit at a decaying rate, spectral betas (Panel (a)). We note that this finding justifies the empirical association in Section 1 between increasing spectral betas and positive first-order delayed adjustments for BRK. We will return to this observation in Section 4.6.

In this model, setting either $\sigma_{u,\xi} = 0$ or $\rho = 0$ eliminates the delayed adjustment and yields constancy of the spectral betas. The case $\sigma_{u,\xi} = 0$ is illustrated in Panel (b). The case $\rho = 0$ is not explicitly represented, but Panels (c) and (d) report the spectral betas for two values of ρ, namely $\rho = 0.1$ and $\rho = 0.8$. The lower persistence, the flatter the spectral betas. As an aside, also, the lower persistence,

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

the more the variance of the regressor (an AR(1), as in the example in Section 4.3) loads, as expected, on lower scales.

We conclude that, in this specification, there is an important interaction between the correlation of the error terms and the persistence of the regressor in leading to both changing spectral betas and delayed adjustments. We will see that the model is close, albeit not identical, to the illustrative specification employed by Kamara et al. (2016) to justify horizon-specific factor loadings upon aggregation (c.f. Section 5.1).

We now consider the same model but with a genuine delayed adjustment in addition to the one induced by endogeneity and persistence of the regressor. Equation (45) now becomes

$$y_t = \left(\beta + \frac{\sigma_{u,\xi}}{\sigma_\xi^2}\right) x_t - \frac{\sigma_{u,\xi}}{\sigma_\xi^2}\rho x_{t-1} + \delta x_{t-1} + \varepsilon_t. \tag{46}$$

As for the Wold representation, we can now express the matrix $\boldsymbol{C}_k$ as follows:

$$\boldsymbol{C}_k = \begin{pmatrix} 0 & \beta\rho + \delta \\ 0 & \rho \end{pmatrix}^k. \tag{47}$$

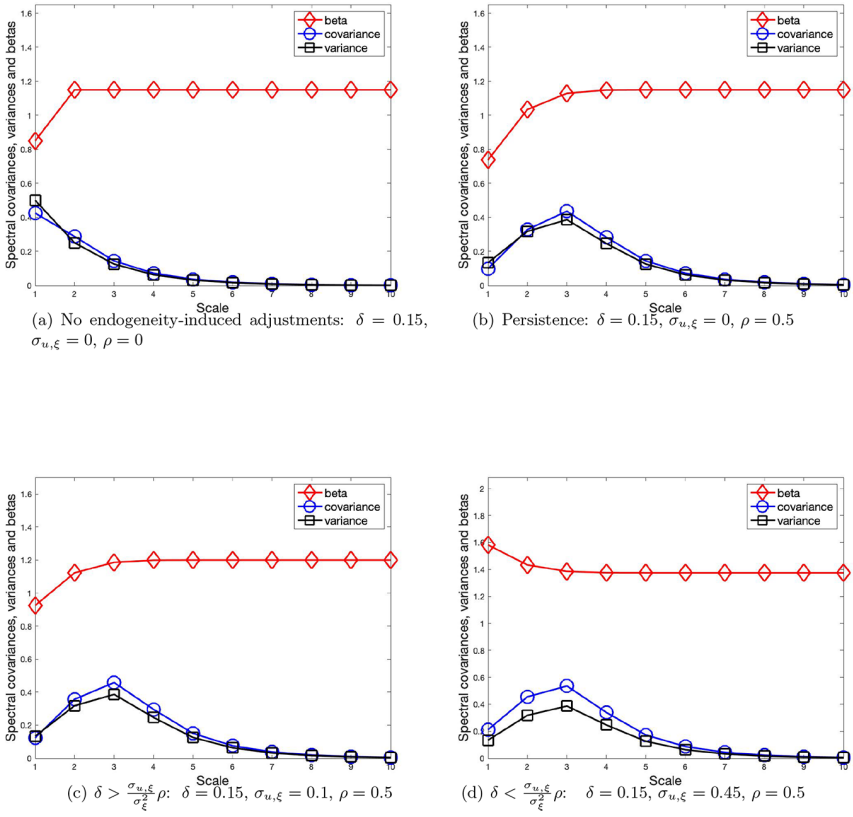
With one genuine delayed adjustment ($\delta \neq 0$), no delayed adjustments due to endogeneity ($\sigma_{u,\xi} = 0$) and no persistence ($\rho = 0$), the first spectral beta is $\beta - \delta$ and all other spectral betas are $\beta + \delta$ (c.f. Remark 16). Figure 4, Panel (a), confirms this result. The case of multiple delayed adjustments is a simple extension with changing betas across scales, to which we are going to return in Section 4.6.

Adding persistence to the regressor ($\rho \neq 0$) would simply lead to a smoother transition between spectral betas (even without lagged adjustments due to endogeneity of the errors, that is, even with $\sigma_{u,\xi} = 0$). Figure 4, Panel (b), provides an illustration.

Finally, in a model with both genuine delayed adjustments (i.e., with $\delta \neq 0$) and delayed adjustments due to endogeneity (i.e., with $\sigma_{u,\xi} \neq 0$), the pattern of the spectral betas should depend on the relation between $\frac{\sigma_{u,\xi}}{\sigma_\xi^2}\rho$ and δ (see equation (46)). Figure 4, Panel (c), considers the case $\delta > \frac{\sigma_{u,\xi}}{\sigma_\xi^2}\rho$. Figure 4, Panel (d), considers the case $\delta < \frac{\sigma_{u,\xi}}{\sigma_\xi^2}\rho$. In the former case, the pattern is increasing, as expected. In the latter, it is decreasing, again, as expected. In both cases, the distance between $\frac{\sigma_{u,\xi}}{\sigma_\xi^2}\rho$ and δ is rather small. A larger distance would, of course, lead to more marked dynamics in the spectral betas. Naturally, when $\delta = \frac{\sigma_{u,\xi}}{\sigma_\xi^2}\rho$, the spectral betas are constant.

We now turn to a two-beta discussion of the link between spectral regression and delayed adjustments to information. The discussion mirrors the derivations leading to equation (27). We employ, however, the extended Wold representation rather than the Haar filter.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press



(a) No endogeneity-induced adjustments: $\delta = 0.15$, $\sigma_{u,\xi} = 0$, $\rho = 0$

(b) Persistence: $\delta = 0.15$, $\sigma_{u,\xi} = 0$, $\rho = 0.5$

(c) $\delta > \frac{\sigma_{u,\xi}}{\sigma_\xi^2}\rho$: $\delta = 0.15$, $\sigma_{u,\xi} = 0.1$, $\rho = 0.5$

(d) $\delta < \frac{\sigma_{u,\xi}}{\sigma_\xi^2}\rho$: $\delta = 0.15$, $\sigma_{u,\xi} = 0.45$, $\rho = 0.5$

FIGURE 4. The spectral betas in a regression model. We represent spectral variances, spectral covariances, and spectral betas for the model in equation (46).

4.5. The Extended Wold: Spectral Regression

Ignoring error terms, as in Section 4.2, assume the process y_t is linked to a "high-frequency" component and a "low-frequency" component of x_t, where x_t is, again, for simplicity but without loss of generality, a scalar process:

$$y_t = \beta_{HF} x_t^{(1)} + \beta_{LF} x_t^{(>1)}. \tag{48}$$

Recall that $x_t^{(>1)} = \sum_{j=2}^{J} x_t^{(j)} + \pi_t^{(J)}$. Given equation (29), we may write

$$x_t^{(1)} = \left(\frac{c_0 - c_1}{\sqrt{2}}\right)\left(\frac{\xi_t - \xi_{t-1}}{\sqrt{2}}\right) + \left(\frac{c_2 - c_3}{\sqrt{2}}\right)\left(\frac{\xi_{t-2} - \xi_{t-3}}{\sqrt{2}}\right) + \cdots,$$

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

then

$$x_t - x_t^{(1)} = \left(\frac{c_0+c_1}{\sqrt{2}}\right)\left(\frac{\xi_t+\xi_{t-1}}{\sqrt{2}}\right) + \left(\frac{c_2+c_3}{\sqrt{2}}\right)\left(\frac{\xi_{t-2}+\xi_{t-3}}{\sqrt{2}}\right) + \cdots$$
$$= \frac{x_{t-1}^{\text{odd}} + x_t^{\text{even}}}{2},$$

with

$$x_t^{\text{even}} = (c_0+c_1)\xi_t + (c_2+c_3)\xi_{t-2} + \cdots$$
$$x_{t-1}^{\text{odd}} = (c_0+c_1)\xi_{t-1} + (c_2+c_3)\xi_{t-3} + \cdots$$

Equation (48), now, gives

$$\begin{aligned} y_t &= \beta_{HF} x_t^{(1)} + \beta_{LF} x_t^{(>1)} \\ &= \beta_{HF} x_t^{(1)} + \beta_{LF}(x_t - x_t^{(1)}) \\ &= \beta_{LF} x_t + (\beta_{HF} - \beta_{LF}) x_t^{(1)} \\ &= \beta_{LF} x_t + (\beta_{HF} - \beta_{LF})\left(x_t - \left(\frac{x_{t-1}^{\text{odd}} + x_t^{\text{even}}}{2}\right)\right) \\ &= \beta_{HF} x_t - \left(\frac{\beta_{HF} - \beta_{LF}}{2}\right)(x_{t-1}^{\text{odd}} + x_t^{\text{even}}) \\ &= \beta_{HF} x_t + \left(\frac{\Lambda_{LF}}{2}\right)(x_{t-1}^{\text{odd}} + x_t^{\text{even}}). \end{aligned} \tag{49}$$

Remark 12. When running a univariate regression of y_t on x_t, the OLS estimate of the parameter associated with x_t is unbiased and consistent (for β_{HF}) only if $\Lambda_{LF} = 0$. In this case, it will—of course—also be unbiased and consistent for $\beta_{LF} = \beta_{HF}$. This observation is intuitive and parallels our discussion following equations (16) and (17).

Remark 13. The term $\left(\frac{\Lambda_{LF}}{2}\right)(x_{t-1}^{\text{odd}} + x_t^{\text{even}})$ is correlated with the regressor x_t unless, of course, $\Lambda_{LF} = 0$. It, therefore, induces endogeneity and "a form of" delayed adjustment (through the dependence on lagged shocks). Coherently with the example in the previous subsection, it is precisely the case $\Lambda_{LF} \neq 0$ which is associated with different spectral betas over alternative frequencies.

Remark 14. The qualifier "a form of" in the previous remark depends on the fact that, even in simple cases, the structure of the adjustment may not be completely transparent. Consider, for example, the AR(1) case in Section 4.3. We have that

$$x_t^{\text{even}} = (1+\rho)(\xi_t + \rho^2\xi_{t-2} + \rho^4\xi_{t-4} + \cdots), \tag{50}$$

and

$$x_{t-1}^{\text{odd}} = (1+\rho)(\xi_{t-1} + \rho^2\xi_{t-3} + \rho^4\xi_{t-5}\cdots). \tag{51}$$

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Now, if we add $(1+\rho)(\rho\xi_{t-2}+\rho^3\xi_{t-4}+\cdots)$ to equation (51) and subtract it from equation (50), equation (51) becomes $(1+\rho)x_{t-1}$ and equation (50) becomes $(1+\rho)(\xi_t+(\rho^2-\rho)\xi_{t-2}+(\rho^4-\rho^3)\xi_{t-4}+\cdots)$, which we call $(1+\rho)\bar{\xi}_t$. Equation (49) can, now, be written as follows:

$$\begin{aligned} y_t &= \beta_{HF}x_t+\left(\frac{\Lambda_{LF}}{2}\right)(x_{t-1}^{\text{odd}}+x_t^{\text{even}}) \\ &= \beta_{HF}x_t+\left(\frac{\Lambda_{LF}(1+\rho)}{2}\right)x_{t-1}+\underbrace{\left(\frac{\Lambda_{LF}(1+\rho)}{2}\right)\bar{\xi}_t}_{u_t}, \end{aligned} \tag{52}$$

which is a delayed adjustment model with endogeneity (unless, again, $\Lambda_{LF}=0$). We note that, if $\rho=0$, the model becomes

$$y_t=\left(\beta_{HF}+\frac{\Lambda_{LF}}{2}\right)\xi_t+\left(\frac{\Lambda_{LF}}{2}\right)\xi_{t-1}, \tag{53}$$

and has a delayed adjustment structure without endogeneity (c.f. equation (27)).

Section 4.4 discussed mechanisms through which delayed adjustment models yield frequency-specific spectral betas (we will expand on empirical implications in Section 4.6). We are now illustrating the converse: models with frequency-specific spectral betas imply forms of delayed adjustment. The latter direction, in particular, is an interesting area for future research.

Remark 15. Because of the orthogonality of the components, the parameters β_{HF} and β_{LF} may be estimated by OLS either jointly, as in equation (48), or in univariate regressions of y_t on the components. Identification of the components hinges on the estimation of the Wold coefficients and shocks. Such identification has been conducted using vector autoregressions (Bandi et al., 2021). Future work should explore the suitability of local projection methods (Jordà, 2005). In both cases, understanding and formalizing the estimation error induced in the slope estimates by the first-stage estimation of the components is of econometric interest.

4.6. Revisiting BRK

We now make use of the conclusions drawn in the previous part of this section to return to BRK. Employing empirical estimates, we simulate from a model in which BRK returns (R_t^{BRK}) depend on contemporaneous market returns (R_t^m) as well as on lagged market returns (R_{t-i}^m):

$$R_t^{\text{BRK}}=\alpha^{\text{BRK}}+\beta R_t^m+\sum_{i=1}^{q}\delta_i R_{t-i}^m+\varepsilon_t. \tag{54}$$

This model is in agreement with the findings in Section 1. There, we reported on an economically sizable (and fairly statistically significant at conventional levels) $\widehat{\delta}_1$ along with four additional lags whose partial effects have a magnitude which is sometimes comparable to that of $\widehat{\delta}_1$.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

The market return follows an AR(1) process:

$$R_t^m = \alpha^m + \rho R_{t-1}^m + \xi_t. \tag{55}$$

Should equations (54) and (55) constitute the right model, ignoring the distributed lag structure in equation (54) would, of course, lead to endogeneity and inconsistent beta estimates. As we discussed in the context of a different, but related, specification in Section 4.4, the lag structure yields frequency-specific spectral betas.

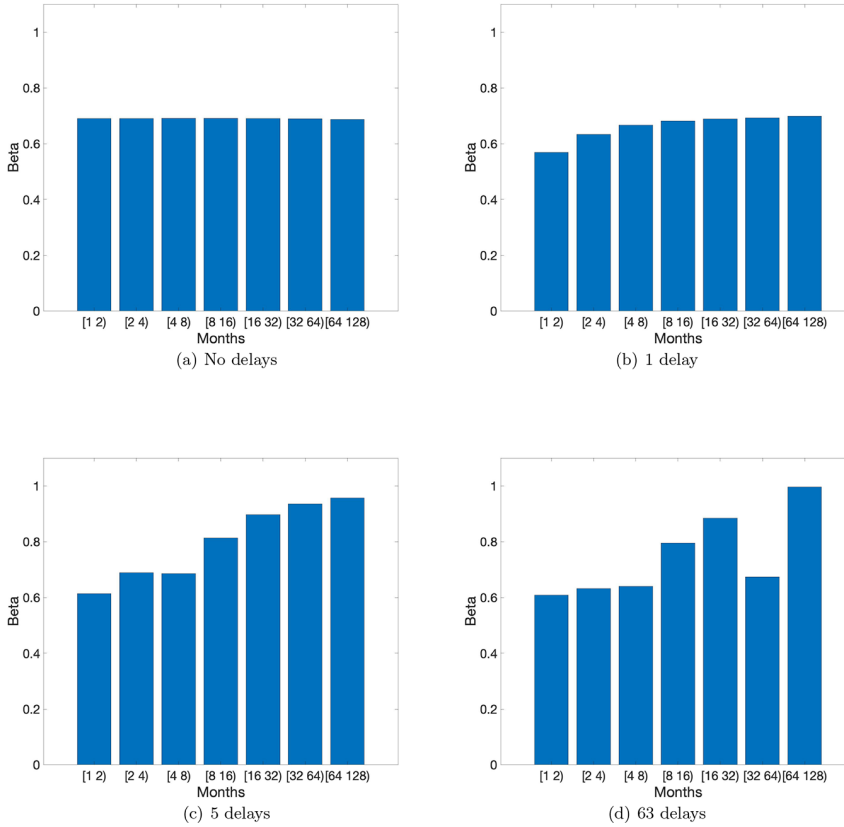
We display the spectral betas obtained by simulating the model without delayed adjustments (i.e., the classical procedure to estimate factor loadings in empirical finance), with one delayed adjustment (i.e., a model inclusive of the only statistically significant—at the 10% level—δ estimate, c.f. Panel (d) of Figure 1), with five delayed adjustments (as in Figure 1, Panels (c) and (d)) and with 63 delayed adjustments (the 63rd being just below the lower end of the horizon spanned by the seventh spectral beta, c.f. Panel (a) of Figure 1). In agreement with the descriptive analysis in Section 1, we use a nonparametric Haar filter to compute the betas from simulated data. The findings are reported in Figure 5.

When $\delta_i = 0$, for all i, the spectral betas are constant. Panel (a) shows that all seven spectral betas are able to yield the overall beta. We recall that the numerical value of the latter is 0.69. With one positive (negative) delayed adjustment, the spectral betas should be increasing (decreasing), albeit at a decreasing rate. Panel (b) confirms this logic, which is also supported by the illustration in Figure 3. We recall that the partial effect (on BRK returns) of the market return delayed by one period is positive (c.f. Panel (c) of Figure 1). Adding four lags to the one-lag model does not modify short-term dynamics but steepens the curve, thereby leading to higher spectral betas over low frequencies (Panel (c)). With 63 lags, we come reasonably close to replicating the pattern of the empirical betas (reported in Panel (a) of Figure 1) with a hump-shape over the first six spectral betas and the largest spectral beta associated with the lowest frequency.

The last result is remarkable. First, it is unlikely that the market is the only driving factor. This said, a model with lagged adjustments to the market *only* appears to perform well at all horizons. Second, and more importantly for the purpose of this discussion, even though we use a model with very long lags which are sometimes large in terms of their magnitude, most of them (with the exception of the first) are statistically insignificant and would, therefore, not be selected by automated selection criteria and used in empirical work. Yet, they result in accurately estimated quantities, that is, the reported spectral betas.

There is a literature on the importance of long adjustments. Campbell and Mankiw (1987), for instance, discuss the limited ability of classical model selection criteria (like AIC or BIC) to choose models with complex lag structures. Relying on the conclusions in Campbell and Mankiw (1987) and simulations, Fuster, Laibson, and Mendel (2010) emphasize the preferability of models with long lags in light of their flexibility in delivering dynamics which more parsimonious specifications would not yield. There is also work on the use of new metrics in the

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press



(a) No delays (b) 1 delay

(c) 5 delays (d) 63 delays

FIGURE 5. Spectral betas from four simulated models. We simulate spectral betas for BRK using a model with no lags, 1 lag, 5 lags, and 63 lags. The parameter estimates used in the simulations are derived from the data described in Section 1.

extraction of signal which classical statistical methods would find hard to identify. The work of Bansal and Yaron (2004) may, for instance, be interpreted as using pricing metrics to extract signal about a persistent component of consumption growth which would otherwise be hard to detect using time-series models applied to consumption data. In the same spirit, the evaluation of frequency-specific effects may be viewed as shining a light, from an alternative vantage point, on models with rich lagged dynamics. We emphasize that frequency-specific metrics (like the spectral betas) have the potential to provide statistically reliable and economically relevant information about the role of shocks in the distant past. We conclude this section with a remark.

Remark 16. Using the Corollary to Theorem 1 of Bandi et al. (2021), it is easy to express the spectral betas for the model specification in equations (54) and (55)

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

in closed form for the case $\rho = 0$. Write $c_0 = \beta$ and $c_k = \delta_k$ for $k > 0$. We have

$$\beta^{(j)} = \sum_{i=0}^{2^{(j-1)}-1} c_i - \sum_{i=0}^{2^{(j-1)}-1} c_{2^{(j-1)}+i} + \frac{\sigma_{u,\xi}}{\sigma_\xi^2}.$$

If the nonzero lagged effects reach the interval $[2^b, 2^{(b+1)})$ with $b \geq 0$, that is, if the maximum lag i such that $\delta_i \neq 0$ is in the interval $[2^b, 2^{(b+1)})$, then the spectral betas are flat from $\beta^{(s)}$ with $s \geq b+2$. In Figure 5, the variability of the spectral betas for $s \geq b+2$ is the result of simulations based on (1) a small (and statistically insignificant) estimated ρ for the market and (2) the use of Haar, rather than Wold-based, components (see Remark 8).

5. AGGREGATION

Aggregation is, arguably, the typical way in which frequency-dependent effects have been detected in empirical finance. In particular, the observation that factor loadings may change depending on the aggregation horizon of the factor (often a market return, like in the BRK example) and an asset return of interest has a long history (c.f. Section 2.2). While this observation has resulted—in the early literature—in searches for the optimal way to *measure* the factor loading of interest, it has now become apparent that there is economic significance to horizon-specific dependence. In other words, horizon-dependent factor loadings may reflect genuine horizon-dependent systematic risk (c.f. Section 2.3).

We begin by linking horizon effects to delayed adjustments in much the same way in which we linked frequency-specific effects to delayed adjustments (Section 5.1). We then turn to the mapping between aggregation and frequency and offer a word of caution (Section 5.2).

5.1. Aggregation and Delayed Adjustments

Assume, for simplicity, that x_t is white noise (a simple subcase of Assumption 1, of course) and scalar. Consider a lead-lag specification of order 1 for y_t:

$$y_t = \beta_1 x_{t+1} + \beta_0 x_t + \beta_{-1} x_{t-1} + \varepsilon_t. \tag{56}$$

Write, now, the covariance between H-period sums of y and x as

$$\sigma_{y,x,H} = \sum_{k=1}^{H} \sum_{l=1}^{H} \mathbb{C}(y_k, x_l). \tag{57}$$

We denote by $\mathbb{C}(y_k, x_k) = \sigma_{y,x,1}(0) = \sigma_{y,x,1}$ the one-period contemporaneous covariances. Under the data generating process in equation (56), the inter-temporal one-period covariances $\mathbb{C}(y_k, x_l) = \sigma_{y,x,1}(k-l)$ for which $|k-l| > 1$ are, of course, zero. Instead, the inter-temporal one-period covariances $\sigma_{y,x,1}(k-l)$ for which $k-l = 1$ are equal to $\rho_{y,x,1}(+1)\sigma_x\sigma_y$ and, as is well-understood, there are $H-1$ of them. Likewise, the $H-1$ intertemporal one-period covariances for which

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

$k-l=-1$ are equal to $\rho_{y,x,1}(-1)\sigma_x\sigma_y$. It follows that

$$\begin{aligned}\sigma_{y,x,H} &= H\sigma_{y,x,1}+(H-1)\rho_{y,x,1}(+1)\sigma_x\sigma_y+(H-1)\rho_{y,x,1}(-1)\sigma_x\sigma_y\\ &= H\sigma_{y,x,1}+(H-1)\left(\rho_{y,x,1}(+1)+\rho_{y,x,1}(-1)\right)\sigma_x\sigma_y\\ &= H\sigma_{y,x,1}+(H-1)\left(\rho_{y,x,1}(+1)+\rho_{y,x,1}(-1)\right)\frac{\sigma_{y,x,1}}{\rho_{y,x,1}}\\ &= \sigma_{y,x,1}\left(H+(H-1)\frac{\rho_{y,x,1}(+1)+\rho_{y,x,1}(-1)}{\rho_{y,x,1}}\right). \end{aligned} \tag{58}$$

This is equation (2) in Hawawini (1983) who assumes that y_t and x_t are continuously compounded returns which implies that the H-period sums are, of course, just H-period logarithmic returns. Under the assumption that x_t is white noise,[18] we also have that the H-period variance of x_t satisfies $\sigma^2_{x,H}=H\sigma^2_x$. It, therefore, follows that

$$\begin{aligned}&\frac{\mathbb{C}\left(\sum_{k=1}^{H}y_k,\sum_{l=1}^{H}x_l\right)}{\mathbb{V}\left(\sum_{k=1}^{H}x_k\right)}\\ &=\frac{\sigma_{y,x,H}}{\sigma^2_{x,H}}\\ &=\frac{\sigma_{y,x,1}\left(H+(H-1)\frac{\rho_{y,x,1}(+1)+\rho_{y,x,1}(-1)}{\rho_{y,x,1}}\right)}{H\sigma^2_x}\\ &=\frac{\sigma_{y,x,1}}{\sigma^2_x}+\left(\frac{H-1}{H}\right)\frac{\sigma_x\sigma_y\left(\rho_{y,x,1}(+1)+\rho_{y,x,1}(-1)\right)}{\sigma^2_x}. \end{aligned} \tag{59}$$

Finally, we may write

$$\beta_H=\beta_0+\left(\frac{H-1}{H}\right)(\beta_1+\beta_{-1}). \tag{60}$$

Equation (60) provides an expression for the H-period beta (β_H) as a function of the one-period contemporaneous and lagged betas (β_0, β_1, and β_{-1}). Its generalization to H leads and lags and, again, white noise x_t observations is natural:

$$\beta_H=\beta_0+\sum_{k=1}^{H-1}\left(\frac{H-k}{H}\right)(\beta_k+\beta_{-k}). \tag{61}$$

Equation (60) is related to the illustrative example represented by equation (6) in Kamara et al. (2016). To see this, set $\beta_1=0$ in equation (56) (no lead effects). As a consequence, equation (60) collapses to

$$\beta_H=\beta_0+\left(\frac{H-1}{H}\right)\frac{\sigma_{y,x,1}(+1)}{\sigma^2_x}. \tag{62}$$

[18] Hawawini (1983) assumes a nonzero first-order autocorrelation for the market, that is, his x_t.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Now, write

$$\begin{aligned} y_t &= \beta_0 x_t + \beta_{-1} x_{t-1} + \varepsilon_t \\ &= \beta_0 x_t + u_t, \end{aligned} \tag{63}$$

where $u_t = \beta_{-1} x_{t-1} + \varepsilon_t$. As in Kamara et al. (2016), we have

$$\mathbb{C}(u_t, x_{t-1}) = \sigma_{y,x,1}(+1) \neq 0.$$

Thus, the error term captures delayed reaction in the simplest (one-period) specification considered by Kamara et al. (2016). The model is, of course, a subcase of the more general specification considered, for example, in equation (54). Kamara et al. (2016) conclude that horizon-dependence in estimated betas upon *aggregation* derives, similarly to the dependence on frequency examined earlier, from slow or delayed adjustments to information.

5.2. Aggregation and Frequency

We begin with the link between aggregation and the frequency-specific components. We then express the factor loadings obtained using aggregates of the original series in terms of factor loadings on the components. In order to facilitate intuition, this discussion is based on frequency-specific components derived from the Haar wavelet reviewed in Section 4.1. As indicated earlier, the structure of the Haar wavelet is particularly helpful to deduce implications for aggregates of time series. Similar implications, albeit less explicitly, apply when mapping time-series aggregates to components in the extended Wold representation, as made clear by the work of Bandi and Tamoni (2021).

Consider 2 processes y_t and x_t and the dyadic horizon of aggregation $H = 2^s$. In their Proposition 2 (Appendix A), Bandi and Tamoni (2021) show that

$$\frac{1}{2^s}\sum_{i=1}^{2^s} y_{t+i} := y_{t,t+2^s} = \sum_{j=s+1}^{J} y^{(j)}_{t+2^s} + \pi^{(y,J)}_{t+2^s}, \tag{64}$$

and

$$\frac{1}{2^s}\sum_{i=1}^{2^s} x_{t+i} := x_{t,t+2^s} = \sum_{j=s+1}^{J} x^{(j)}_{t+2^s} + \pi^{(x,J)}_{t+2^s}, \tag{65}$$

where $s = 0, 1, \ldots, J-1$ denotes the aggregation level and $\left\{\pi^{(y,J)}_{t+2^s}, \pi^{(x,J)}_{t+2^s}\right\}$ are long-run averages defined in equation (21). Aggregation over a specific horizon reveals information over frequencies *lower* than the one corresponding to the aggregation level.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

We now turn to the regression slopes. In their Proposition 3 (Appendix A), Bandi and Tamoni (2021) show that, as $J \to \infty$,

$$\beta_s = \frac{\mathbb{C}\left(y_{t,t+2^s}, x_{t,t+2^s}\right)}{\mathbb{V}\left(x_{t,t+2^s}\right)} = \sum_{j=s+1}^{\infty} v^{(j)} \beta^{(j)}, \tag{66}$$

which is an extension of equation (37). The regression slopes with respect to aggregates of the original series can be expressed as weighted averages of regression slopes with respect to lower (than the aggregation horizon) frequency components. Because averaging filters out transient dynamics, more averaging uncovers lower frequency components. Finally, it is informative to re-write equation (66) as follows:

$$\beta_s = \underbrace{v^{(s+1)} \beta^{(s+1)}}_{\text{signal}} + \underbrace{\sum_{j=s+2}^{\infty} v^{(j)} \beta^{(j)}}_{\text{noise}}, \tag{67}$$

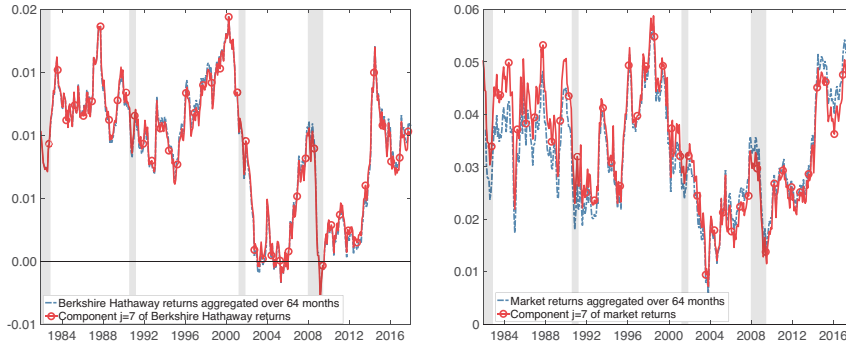
where $v^{(j)} = \frac{\mathbb{V}(x^{(j)})}{\mathbb{V}(x_{t,t+2^s})}$, with $j \geq s+1$.

Bandi and Tamoni (2021) warn against the indiscriminate use of aggregation as a way to analyze frequency-specific effects. If the interest is in very low-frequency quantities, aggregation is likely to yield reliable information. Mathematically, the larger the aggregation level s, the smaller the noise component in equation (67), the closer the ratio $v^{(s+1)}$ to 1 and, importantly, the closer β_s to $\beta^{(s+1)}$. Should interest be in higher-frequency fluctuations, then both the noise component and the weight would contribute in creating a separation between betas upon aggregation and betas on the components. While one could bias correct (to eliminate both the noise component and the weight), the bias-correction would require computation of the components. Availability of the components would, however, defeat the purpose of aggregating: if interest is in frequency-specific dynamics, computing the frequency-specific components is, in general, a more robust strategy than aggregating.

Next, we return to our leading example and illustrate a case in which aggregation works very satisfactorily as a method for extracting frequency-specific information. We display the component with fluctuations between 64 and 128 months of the BRK returns (i.e., $y^{(j)}_{t+2^s}$ with $j = 7$) against the BRK returns aggregated over 64 months (i.e., $y_{t,t+2^s}$). In light of equation (64), the two should be similar if the residual component is small in terms of its variance (and we find it to be small). We do the same for the market returns. Panels (a) and (b) of Figure 6 provide a visualization. Frequency-specific components and aggregates are very close to each other. Not surprisingly, when proxying for the low-frequency spectral beta of BRK using beta on aggregates (i.e., β_s) we find a value (1.457) which is virtually identical to the value reported in Panel (a) of Figure 1.

The results in Bandi and Tamoni (2021) illustrate how a similar logic applies to the aggregation of covariance-stationary time series, that is, time series which

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press



(a) BRK (b) The market

FIGURE 6. Aggregates versus frequency-specific components. The left panel displays the BRK component with fluctuations between 64 and 128 months against the BRK returns aggregated over 64 months. The right panel provides the same information for the market. The gray shaded lines are NBER recessions. We use the data described in Section 1.

can be expressed in terms of a Wold representation and whose frequency-specific components are described by the components of the extended Wold representation. For these time series, we show that the *aggregated* process also satisfies an extended Wold representation for which the components can easily be quantified in terms of the Wold parameters of the original series. As in the case of the Haar filter, aggregation (over an horizon s, say) reduces the influence (i.e., the variance) of components with $j \leq s$ (given j, the larger s, the larger the variance reduction). These components do not vanish (as in the Haar case illustrated in equations (64) and (65)) but their variance is drastically attenuated. As a result, one is left with an aggregated process which loads on low-frequency components (with $j > s$) that are closely related to the low-frequency components of the disaggregated process. As is the case for the Haar filter (c.f., again, equations (64) and (65)), the aggregation of covariance-stationary time series can, therefore, be viewed as a mechanism to extract their low-frequency cycles. Bandi and Tamoni (2021) provide details and closed-form formulations using models with *implicit* and *explicit* delayed adjustments, as in Section 4.4.

6. CONCLUSIONS AND FURTHER DIRECTIONS FOR FUTURE WORK

The goal of this article was to review methods of statistical inference suited to provide information about interactions between economic cycles of different length. To this end, we have reviewed spectrum regression, a tool explicitly designed to localize on frequency. We have mapped spectrum regression to recent developments in multiresolution analysis. Because of its traditional emphasis on both time and scales and its recent emphasis on statistical validation, this is an area which has made considerable strides in bridging the gap between time domain

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

and frequency domain, thereby facilitating economic interpretation as a result. We have also linked it to aggregation, arguably the conventional way in which horizon-specific effects have been captured, at least in finance.

We illustrated, by way of a formal discussion and a leading empirical example, the economic consequences of rich dependence on frequency. In general, frequency-dependent relations translate into possibly nontrivial lead-lag structures. When understood thoroughly, such structures may contain a good deal of information about the timing and intensity of economic shocks. For instance, the existing evidence on frequency-dependent systematic risk in financial markets is consistent with a delay in the reaction of asset prices to news about fundamental sources of risk. Both short and long delays matter, in general, thereby raising questions about the mechanisms through which news leads to valuations and, relatedly, about the nature of market efficiency.

Much remains to be done. On the empirical side, a closer look at dynamics over alternative frequencies, scales or horizons has the potential to be informative about reactions to shocks in the very distant past which would otherwise be hard to detect using conventional methods of inference (c.f. Section 4.6). The result may be more adequate modeling in the time domain with favorable implications in terms of enhanced economic interpretability and, hopefully, superior forecasting ability.

Regarding economic interpretability, there has been a tendency—at least in finance—to aggregate as a way to handle transient effects and deal with measurement error (c.f. Section 2.2). In essence, an implicit assumption has often been made that, because economic theory is silent about the way in which we should think about time, evaluating economic relations over different horizons is an opportune way to assess the validity of economic models while handling the inherent noise that the data contains, especially at higher frequencies. This approach is understandable and has led to numerous insights. However, it is arguably only a first step, one which is divorced from attention to the role of dynamics over alternative frequencies/horizons, their consistency, and their implications for time series modeling and economic reasoning (c.f. Section 2.3).

Returning to our leading example, there is a fundamental economic difference between stating that an alleged risk factor, like the market, is better validated by data (as a *priced* risk factor) when returns are aggregated over longer periods and saying that *low-frequency* market fluctuations appear to be valid risk factors when pricing *higher-frequency* asset returns. In the former case, the classical approach in the literature, one is simply being agnostic about the horizon over which to test a pricing model and aggregation is used as a low-pass filter to handle transient (unpriced) dynamics. In the latter case, dynamics over alternative frequencies have a genuine role to play and pricing itself is symptomatic of the interaction between frequencies and the need for coherent economic logic.

In terms of economic modeling, the former is simply relying on classical economic theory and the relabeling of time and horizon (from months, to quarters,

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

to years). The latter is asking for a novel framework in which heterogeneity in decision making over horizons, horizon-specific risk aversion and the like are not just anecdotal but are, in fact, convincingly formalized.

On the econometrics side, while frequency-domain methods are well-developed, the application of multiresolution techniques to economic shocks, rather than to data directly (in the form of a nonparametric filter), has brought attention to signal extraction techniques as a way to model economic time series (c.f. Sections 4.3 and 4.4). Now that multiresolution methods can be viewed through the lens of traditional time-series specifications, the rich apparatus of traditional time series analysis would apply to them, thereby opening up new possibilities.

For all of its recent emphasis on the long run, surprisingly little work in finance is currently stepping outside of the boundaries of time-domain specifications. We have argued that moving outside of those boundaries may, in fact, help model selection in the time domain and assist, rather than impair, economic logic.

REFERENCES

Andersen, T.G. & R.T. Varneskov (2021) Consistent inference for predictive regressions in persistent economic systems. *Journal of Econometrics* 224, 215–244.

Bandi, F.M., S. Chaudhuri, A.W. Lo, & A. Tamoni (2021) Spectral factor models. *Journal of Financial Economics* 142, 214–238.

Bandi, F.M. & B. Perron (2006) Long memory and the relation between implied and realized volatility. *Journal of Financial Econometrics* 4, 636–670.

Bandi, F.M., B. Perron, A. Tamoni, & C. Tebaldi (2019) The scale of predictability. *Journal of Econometrics* 208, 120–140.

Bandi, F.M. & Y. Su (2022) Conditional Spectral Methods. Working paper, Johns Hopkins University.

Bandi, F.M. & A. Tamoni (2021) Business-Cycle Consumption Risk and Asset Prices. Working paper, Johns Hopkins University and Rutgers University.

Bansal, R. & A. Yaron (2004) Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59, 1481–1509.

Baxter, M. & R.G. King (1999) Measuring business cycles: Approximate band-pass filters for economic time series. *Review of Economics and Statistics* 81, 575–593.

Berkowitz, J. (2001) Generalized spectral estimation of the consumption-based asset pricing model. *Journal of Econometrics* 104, 269–288.

Beveridge, S. & C.R. Nelson (1981) A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the "business cycle". *Journal of Monetary Economics* 7, 151–174.

Burns, A.F. & W.C. Mitchell (1946) *Measuring Business Cycles*. National Bureau of Economic Research.

Campbell, J.Y. & G. Mankiw (1987) Are output fluctuations transitory? *Quarterly Journal of Economics* 102, 857–880.

Chaudhuri, S.E. & A.W. Lo (2015) Spectral analysis of stock-return volatility, correlation, and beta. In *2015 IEEE Signal Processing and Signal Processing Education Workshop*, pp. 232–236.

Chaudhuri, S.E. & A.W. Lo (2018) Dynamic alpha: A spectral decomposition of investment performance across time horizons. *Management Science* 65, 4440–4450.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Choi, I. & P.C.B. Phillips (1993) Testing for a unit root by frequency domain regression. *Journal of Econometrics* 59, 263–286.
Christensen, B.J. & M.Ø. Nielsen (2006) Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting. *Journal of Econometrics* 133, 343–371.
Christiano, L.J. & T.J. Fitzgerald (2003) The band pass filter. *International Economic Review* 44, 435–465.
Christiano, L.J. & R.J. Vigfusson (2003) Maximum likelihood in the frequency domain: The importance of time-to-plan. *Journal of Monetary Economics* 50, 789–815.
Cogley, T. (2001) A frequency decomposition of approximation errors in stochastic discount factor models. *International Economic Review* 42, 473–503.
Cohen, L. & A. Frazzini (2008) Economic links and predictable returns. *Journal of Finance* 63, 1977–2011.
Cohen, K.J., G.A. Hawawini, S.F. Maier, R.A. Schwartz, & D.K. Whitcomb (1983) Friction in the trading process and the estimation of systematic risk. *Journal of Financial Economics* 12, 263–278.
Cohen, R.B., C. Polk, & T. Vuolteenaho (2009) The price is (almost) right. *Journal of Finance* 64, 2739–2782.
Comin, D. & M. Gertler (2006) Medium-term business cycles. *American Economic Review* 96, 523–551.
Corbae, D., S. Ouliaris, & P.C.B. Phillips (1994) A reexamination of the consumption function using frequency domain regressions. *Empirical Economics* 19, 595–609.
Corbae, D., S. Ouliaris, & P.C.B. Phillips (2002) Band spectral regression with trending data. *Econometrica* 70, 1067–1109.
Corsi, F. (2009) A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7, 174–196.
Croux, C., M. Forni, & L. Reichlin (2001) A measure of comovement for economic variables: Theory and empirics. *Review of Economics and Statistics* 83, 232–241.
Daubechies, I. (1992) *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics.
D. Bondt, F.M. Werner, & R. Thaler (1985) Does the stock market overreact? *Journal of Finance* 40, 793–805.
Dew-Becker, I. & S. Giglio (2016) Asset pricing in the frequency domain: Theory and empirics. *Review of Financial Studies* 29, 2029–2068.
Diebold, F.X., L.E. Ohanian, & J. Berkowitz (1998) Dynamic equilibrium economies: A framework for comparing models and data. *Review of Economic Studies* 65, 433–451.
Dimson, E. (1979) Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics* 7, 197–226.
Engle, R.F. (1974) Band spectrum regression. *International Economic Review* 15, 1–11.
Engle, R.F. (1978) Testing price equations for stability across spectral frequency bands. *Econometrica* 46, 869–881.
Englund, P., T. Persson, & L.E.O. Svensson (1992) Swedish business cycles: 1861–1988. *Journal of Monetary Economics* 30, 343–371.
Fama, E.F. & K.R. French (1992) The cross-section of expected stock returns. *Journal of Finance* 47, 427–465.
Faria, G. & F. Verona (2018) Forecasting stock market returns by summing the frequency-decomposed parts. *Journal of Empirical Finance* 45, 228–242.
Faria, G. & F. Verona (2020) Time-frequency forecast of the equity premium. *Quantitative Finance* 21, 1–17.
Feng, G., S. Giglio, & D. Xiu (2020) Taming the factor zoo: A test of new factors. *Journal of Finance* 75, 1327–1370.
Freyberger, J., A. Neuhierl, & M. Weber (2020) Dissecting characteristics nonparametrically. *Review of Financial Studies* 33, 2326–2377.
Friedman, M. (1957) *A Theory of the Consumption Function*. Princeton University Press.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Friedman, M. & S. Kuznets (1954) *Income from Independent Professional Practice*. National Bureau of Economic Research.

Fuster, A., D. Laibson, & B. Mendel (2010) Natural expectations and macroeconomic fluctuations. *Journal of Economic Perspectives* 24, 67–84.

Gençay, R., F. Selçuk, & B.J. Whitcher (2001) *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Elsevier.

Gençay, R., F. Selçuk, & B.J. Whitcher (2003) Systematic risk and timescales. *Quantitative Finance* 3, 108–116.

Gençay, R., F. Selçuk, & B.J. Whitcher (2005) Multiscale systematic risk. *Journal of International Money and Finance* 24, 55–70.

Gilbert, T., C. Hrdlicka, J. Kalodimos, & S. Siegel (2014) Daily data is bad for beta: Opacity and frequency-dependent betas. *Review of Asset Pricing Studies*, 4, 78–117.

Granger, C.W.J. & J.-L. Lin (1995) Causality in the long run. *Econometric Theory* 11, 530–536.

Granger, C.W.J. & M.W. Watson (1984) Time series and spectral methods in econometrics. In Z. Griliches & M.D. Intriligator (eds). *Handbook of Econometrics*, vol. 2, pp. 979–1022. Elsevier.

Haar, A. (1910) Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen* 69, 331–371.

Hamilton, J.D. (2018) Why you should never use the Hodrick-Prescott filter. *Review of Economics and Statistics* 100, 831–843.

Handa, P., S.P. Kothari, & C. Wasley (1989) The relation between the return interval and betas: Implications for the size effect. *Journal of Financial Economics* 23, 79–100.

Handa, P., S.P. Kothari, & C. Wasley (1993) Sensitivity of multivariate tests of the capital asset-pricing model to the return measurement interval. *Journal of Finance* 48, 1543–1551.

Hannan, E.J. (1963a) Regression for time series. In M. Rosenblatt (ed). *Time Series Analysis*, pp. 17–37. Wiley.

Hannan, E.J. (1963b) Regression for time series with errors of measurement. *Biometrika* 50, 293–302.

Hannan, E.J. (1969) The estimation of mixed moving average autoregressive systems. *Biometrika* 56, 579–593.

Hannan, E.J. (1970) *Multiple Time Series*. Wiley.

Hannan, E.J. & P.M. Robinson (1973) Lagged regression with unknown lags. *Journal of the Royal Statistical Society, Series B* 35, 252–267.

Hannan, E.J. & P.J. Thomson (1971) Spectral inference over narrow bands. *Journal of Applied Probability* 8, 157–169.

Harvey, A.C. (1978) Linear regression in the frequency domain. *International Economic Review* 19, 507–512.

Harvey, C.R., Y. Liu, & H. Zhu (2016) . . . and the cross-section of expected returns. *Review of financial studies* 29, 5–68.

Hawawini, G. (1983) Why beta shifts as the return interval changes. *Financial Analysts Journal* 39, 73–77.

Heston, S.L. & R. Sadka (2008) Seasonality in the cross-section of stock returns. *Journal of Financial Economics* 87, 418–445.

Hodrick, R.J. (2020) An Exploration of Trend-Cycle Decomposition Methodologies in Simulated Data. Working paper, Columbia University.

Hodrick, R.J. & E.C. Prescott (1997) Postwar US business cycles: An empirical investigation. *Journal of Money, Credit, and Banking* 29, 1–16.

Hong, H. & J.C. Stein (1999) A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance* 54, 2143–2184.

Hong, H., W. Torous, & R. Valkanov (2007) Do industries lead stock markets? *Journal of Financial Economics* 83, 367–396.

Hosoya, Y. (1991) The decomposition and measurement of the interdependency between second-order stationary processes. *Probability Theory and Related Fields* 88, 429–444.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Hou, K. & T.J. Moskowitz (2005) Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies* 18, 981–1020.
Jegadeesh, N. (1990) Evidence of predictable behavior of security returns. *Journal of Finance* 45, 881–898.
Jordà, Ò. (2005) Estimation and inference of impulse responses by local projections. *American Economic Review* 95, 161–182.
Kamara, A., R.A. Korajczyk, X. Lou, & R. Sadka (2016) Horizon pricing. *Journal of Financial and Quantitative Analysis* 51, 1769–1793.
Kang, B.U., F. In, & T.S. Kim (2017) Timescale betas and the cross section of equity returns: Framework, application, and implications for interpreting the Fama–French factors. *Journal of Empirical Finance* 42, 15–39.
Keim, D.B. (1983) Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of Financial Economics* 12, 13–32.
Kelly, B.T., S. Pruitt, & Y. Su (2019) Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics* 134, 501–524.
Kim, S., R.A. Korajczyk, & A. Neuhierl (2021) Arbitrage portfolios. *Review of Financial Studies* 34, 2813–2856.
King, R.G. & M.W. Watson (1996) Money, prices, interest rates and the business cycle. *Review of Economics and Statistics* 78, 35–53.
Kothari, S.P., J. Shanken, & R.G. Sloan (1995) Another Look at the Cross-Section of Expected Stock Returns. *Journal of Finance* 50, 185–224.
Kozak, S., S. Nagel, & S. Santosh (2020) Shrinking the cross-section. *Journal of Financial Economics* 135, 271–292.
Lettau, M. & M. Pelger (2020) Factors that fit the time series and cross-section of stock returns. *Review of Financial Studies* 33, 2274–2325.
Levhari, D. & H. Levy (1977) The capital asset pricing model and the investment horizon. *Review of Economics and Statistics* 59, 92–104.
Lintner, J. (1965) The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47, 13–37.
Lo, A.W. & A. Craig MacKinlay (1990) An econometric analysis of nonsynchronous trading. *Journal of Econometrics* 45, 181–211.
Mallat, S.G. (1989) A Theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11, 674–693.
Marinucci, D. (2000) Spectral regression for cointegrated time series with long-memory innovations. *Journal of Time Series Analysis* 21, 685–705.
Marinucci, D. & P.M. Robinson (2001) Semiparametric fractional cointegration analysis. *Journal of Econometrics* 105, 225–247.
Marinucci, D. & P.M. Robinson (2003) Semiparametric frequency domain analysis of fractional cointegration. In P.M. Robinson (ed.), *Time Series with Long Memory*, pp. 334–371. Oxford University Press.
Markowitz, H. (1952) Portfolio selection. *Journal of Finance* 7, 77–91.
Müller, U.K. & M.W. Watson (2008) Testing models of low-frequency variability. *Econometrica* 76, 979–1016.
Müller, U.K. & M.W. Watson (2016) Measuring uncertainty about long-run predictions. *Review of Economic Studies* 83, 1711–1740.
Müller, U.K. & M.W. Watson (2017) Low-Frequency Econometrics. In B. Honoré, A. Pakes, M. Piazzesi, & L. Samuelson (eds.), *Advances in Economics and Econometrics*, vol. 2, pp. 53–94. Cambridge University Press.
Müller, U.K. & M.W. Watson (2018) Long-run covariability. *Econometrica* 86, 775–804.
Neuhierl, A. & R.T. Varneskov (2021) Frequency dependent risk. *Journal of Financial Economics* 140, 644–675.
Ortu, F., F. Severino, A. Tamoni, & C. Tebaldi (2020) A persistence-based Wold-type decomposition for stationary time series. *Quantitative Economics* 11, 203–230.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press

Ortu, F., A. Tamoni, & C. Tebaldi (2013) Long-run risk and the persistence of consumption shocks. *Review of Financial Studies* 26, 2876–2915.

Otrok, C., B. Ravikumar, & C.H. Whiteman (2002) Habit formation: A resolution of the equity premium puzzle? *Journal of Monetary Economics* 49, 1261–1288.

Phillips, P.C.B. (1991) *Spectral regression for cointegrated time series*. In W. Barnett, J. Powell, & G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Economics and Statistics*, pp. 413–435. Cambridge University Press.

Phillips, P.C.B. (1998) New tools for understanding spurious regressions. *Econometrica* 66, 1299–1325.

Phillips, P.C.B. (2005) HAC estimation by automated regression. *Econometric Theory* 21, 116–142.

Phillips, P.C.B. & S. Jin (2020) Business cycles, trend elimination, and the HP filter. *International Economic Review* 62, 469–520.

Phillips, P.C.B. & M. Loretan (1991) Estimating long-run economic equilibria. *Review of Economic Studies* 58, 407–436.

Phillips, P.C.B. & Z. Shi (2020) Boosting: Why you can use the HP filter. *International Economic Review* 62, 521–570.

Robinson, P.M. (1977) The construction and estimation of continuous-time models and discrete approximations in econometrics. *Journal of Econometrics* 6, 173–197.

Roll, R. (1977) A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. *Journal of Financial Economics* 4, 129–176.

Rozeff, M.S. & W.R. Kinney (1976) Capital market seasonality: The case of stock returns. *Journal of Financial Economics* 3, 379–402.

Sarno, L., D.L. Thornton, & Y. Wen (2007) What's unique about the federal funds rate? Evidence from a spectral perspective. *Oxford Bulletin of Economics and Statistics* 69, 293–319.

Scholes, M. & J. Williams (1977) Estimating betas from nonsynchronous data. *Journal of Financial Economics* 5, 309–327.

Schwartz, R.A. & D.K. Whitcomb (1977) Evidence on the presence and causes of serial correlation in market model residuals. *Journal of Financial and Quantitative Analysis* 12, 291–313.

Sharpe, W.F. (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19, 425–442.

Shensa, M.J. (1992) The discrete wavelet transform: Wedding the à trous and Mallat algorithms. *IEEE Transactions on Signal Processing* 40, 2464–2482.

Smith, K.V. (1978) The effect of intervaling on estimating parameters of the capital asset pricing model. *Journal of Financial and Quantitative Analysis* 13, 313–332.

Thomson, P.J. (1986) Band-limited spectral estimation of autoregressive-moving-average processes. *Journal of Applied Probability* 23, 143–155.

Watson, M.W. (1993) Measures of fit for calibrated models. *Journal of Political Economy* 101, 1011–41.

Xiao, Z. & P.C.B. Phillips (1998) Higher-order approximations for frequency domain time series regression. *Journal of Econometrics* 86, 297–336.

Yu, J. (2012) Using long-run consumption-return correlations to test asset pricing models. *Review of Economic Dynamics* 15, 317–335.

https://doi.org/10.1017/S0266466622000020 Published online by Cambridge University Press