MIDAS Regressions and Their Applications in Finance and Macroeconomics

Eric Ghysels

SAMSI Lecture notes - September 2005

Introduction

- The idea to construct regressions combining data with different sampling frequencies is explored. Think of combining annual and quarterly/monthly data, monthly/daily, daily/intradaily, etc.

- We call the regression framework a MIxed DAta S$ampling$ regression (MIDAS regression).

- Suppose y_t is sampled at some fixed, say annual, quarterly, monthly or daily, frequency called interval of reference.

- Denote by $x_t^{(m)}$ a process sampled m times during interval of reference.

- We can write a simple linear MIDAS regression:

$$y_t = \beta_0 + \sum_{j=1}^{j^{max}} b(j,\theta) x_{t-j/m}^{(m)} + \varepsilon_t = \beta_0 + B(L^{1/m}) x_t^{(m)} + \varepsilon_t$$

Where $B(L^{1/m}) = b(0,\theta) + b(1,\theta)L^{1/m} + \ldots + b(j^{max},\theta)L^{j^{max}/m}$ is a polynomial of length j^{max} governed by small set of hyperparameters θ, and $L^{j/m}x_t^{(m)} = x_{t-j/m}^{(m)}$, such that $(L^{1/m})^m = L$

- Relates to distributed lag models

$$y_{t+1} = \beta_0 + \sum_{j=0}^{j^{max}} b(j,\theta) x_{t-j} + \varepsilon_{t+1} = \beta_0 + B(L)x_t + \varepsilon_{t+1}$$

where $B(L)$ is some finite or infinite lag polynomial operator, usually parameterized by a small set of hyperparameters θ.

- See e.g. Dhrymes (1971) and Sims (1974) for surveys on distributed lag models. Many econometrics textbooks also cover the topic, see e.g. Greene (2000, chap. 17), Judge et al. (1985, chap. 9 - 10), Stock and Watson (2003, chap. 13) Wooldridge (2000, chap. 18), among others.

Some material is taken from papers downloadable at:

`http://www.unc.edu/~eghysels/`

- *There is a Risk-Return Tradeoff After All* (*JFE,* forthcoming)
- *The MIDAS Touch: Mixed Data Sampling Regression Models*
- *Predicting Volatility: Getting the Most out of Data Sampled at Different Frequencies*(*JEconometrics,* forthcoming)

All of the above with P. Santa-Clara and R. Valkanov

- *MIDAS Regressions: Further Results and New Directions*, with P. Santa-Clara, A. Sinko and R. Valkanov

- *Why Is Realized Absolute Value Such A Good Predictor Of Volatility?*, with L. Forsberg

- *The impact of economic news on the cross section of returns*, with A. Sinko and R. Valkanov

- *Sales, Promotion and Financial Performance*, with K. Pauwels

- *Forecasting Professional Forecasters*, with J. Wright

- *Semiparametric MIDAS regressions*, with E. Renault

Motivating Examples

Example I: Risk-Return Trade-off

- The risk-return tradeoff involves the following regression:

$$R_{t+1} = \mu + \gamma \widehat{\sigma}_t^2 + \epsilon_{t+1} \tag{1}$$

where R_{t+1} is the excess return on the market in month $t+1$, and $\widehat{\sigma}_t^2$ is the forecasted variance of returns for the same month $t+1$, based on information known at time t.

- French et al. (1987) use within-month daily returns to estimate the realized variance in the period from $t-1$ to t (where typically $D = 22$, 44, 66, etc.):

$$\widehat{\sigma}_t^2 = \sum_{j=1}^{D} r_{t-j/22}^2 \tag{2}$$

Results with CRSP VW excess returns - Jan. 1946 to Dec. 2000

Window (Months)	Conditional Mean Equation μ	γ	R^2
1	0.0107	-0.3422	0.0004
	(5.6932)	(-0.5365)	
2	0.0085	1.2330	0.0034
	(4.2150)	(1.5041)	
3	0.0073	2.0328	0.0072
	(3.4309)	(2.1725)	
12	0.0085	1.4310	0.0015
	(3.1820)	(0.9704)	

Results with CRSP VW excess returns - With Control Variables

Months

	μ	γ	DP	TSPR	RTBL	R^2
1	0.0105	-0.2145	0.0029	-0.0000	-0.0039	0.0139
	(5.6112)	(-0.3362)	(1.7517)	(-0.0051)	(-2.3851)	
2	0.0084	1.3316	0.0030	-0.0005	-0.0039	0.0178
	(4.1538)	(1.6212)	(1.8502)	(-0.2895)	(-2.4073)	
3	0.0072	2.1267	0.0031	-0.0005	-0.0038	0.0215
	(3.3695)	(2.2628)	(1.9177)	(-0.3196)	(-2.3454)	
12	0.0081	1.7147	0.0031	-0.0010	-0.0039	0.0164
	(2.9925)	(1.1371)	(1.8620)	(-0.5861)	(-2.3855)	

Table reports the same regression with three commonly used forecasters of excess returns: the dividend yield, the term spread between a 10 year bond and the three-month Treasury bill, and the stochastically detrended three-month Treasury bill rate (Campbell (1991)).

Example II: Predicting Realized Volatility

- Andersen, Bollerslev, Diebold and co-authors (2001 a,b, 2002), Andreou and Ghysels (2002), Barndorff-Nielsen and Shephard (2001, 2002 a,b, 2003), Taylor and Xu (1997), model realized volatility, $\tilde{Q}_{t+1}^{(m)}$, based on m intradaily returns:

$$\tilde{Q}_{t+1}^{(m)} = \beta_0 + B(L)\tilde{Q}_t^{(m)} + \varepsilon_{t+1}$$

- Such models are estimated with 'unconstrained' polynomials

- Note again that one first computes $\tilde{Q}_t^{(m)}$ (aggregation) and then runs regression involving aggregated data.

Example III: News impact on the stock market

- The impact of macro and corporate news (low frequency event) on the entire cross section of individual stock returns (high frequency). This involves projecting low frequency data onto high frequency data. There is a substantial literature on the topic. MIDAS regression allows us to explore this further.

Example IV: Forecasting professional forecasters

- MIDAS methodology ideally suited to using high frequency financial data to predict low frequency macro data. One example is to 'forecast professional forecasters' (quarterly or monthly) using financial market data (daily or intra-daily).

Parameterizations of the $B\left(L^{1/m}\right)$ Polynomial

Exponential Almon and Beta Polynomials

- We propose two parameterizations of $b(k;\theta)$.

- The first one is:

$$b(k;\theta) = \frac{e^{\theta_1 k + ... + \theta_Q k^Q}}{\sum_{k=1}^{k^{max}} e^{\theta_1 k + ... + \theta_Q k^Q}} \tag{3}$$

which we call the "Exponential Almon lag," since it is related to "Almon lags" (see e.g. Judge et al. 1985).

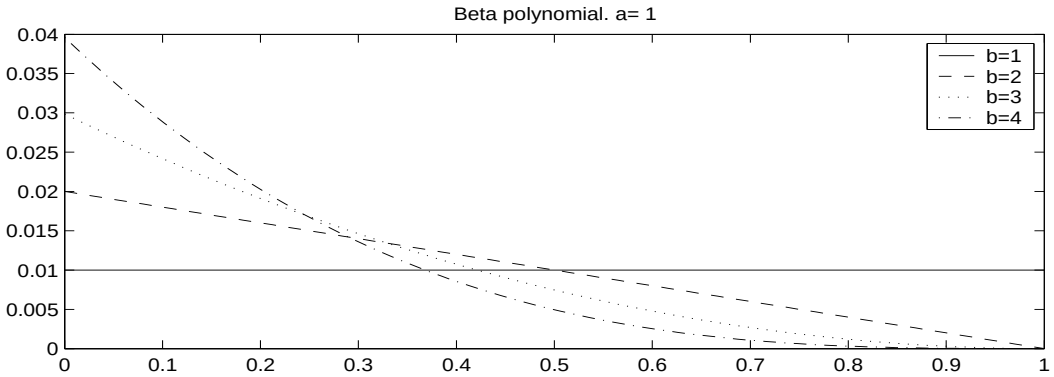
- The second parameterization has only two parameters, or $\theta = [\theta_1; \theta_2]$:

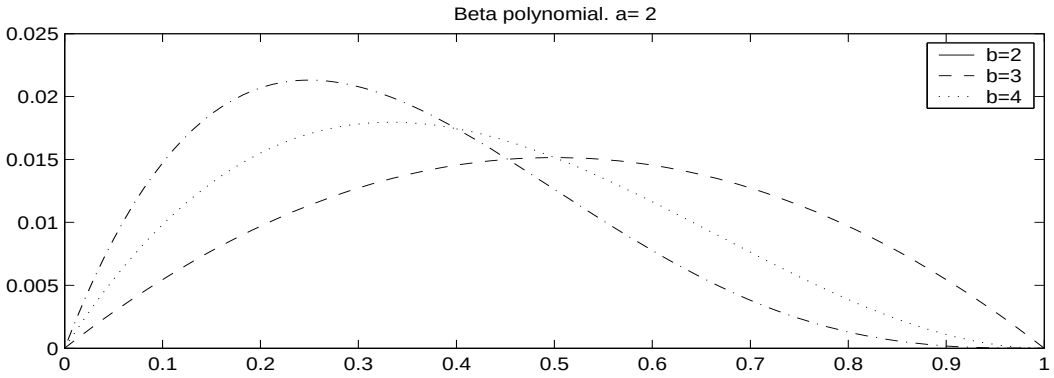
$$b(k; \theta_1, \theta_2) = \frac{f(\frac{j}{j^{max}}, \theta_1; \theta_2)}{\sum_{j=1}^{j^{max}} f(\frac{j}{j^{max}}, \theta_1; \theta_2)} \tag{4}$$

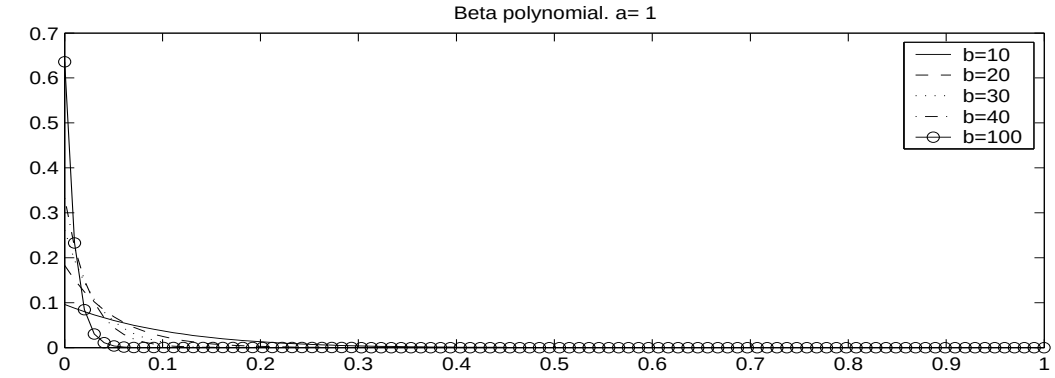
where:

$$\begin{aligned} f(x, a, b) &= \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} \\ B(a, b) &= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \\ \Gamma(a) &= \int_0^\infty e^{-x} x^{a-1} dx \end{aligned}$$

- Specification (**??**) has, to the best of our knowledge, not been used in the literature. It is based on the beta function and we refer to it as the "Beta lag."


Beta polynomial. a= 1
b=1
b=2
b=3
b=4
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1


Beta polynomial. a= 2
b=2
b=3
b=4
0.025
0.02
0.015
0.01
0.005
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1


Beta polynomial. a= 1
b=10
b=20
b=30
b=40
b=100
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

MIDAS with stepfunctions

- MIDAS with stepfunctions (special case is HAR (Heterogenous Autoregressive) Model (Corsi(2003)))

$$\tilde{Q}_{t+1}^{(m)} = \beta_0 + \beta_D X_{t-1,t} + \beta_W X_{t-5,t} + \beta_M X_{t-20,t} + \varepsilon_{t+1}$$

for X various regressors discussed later. The advantage of using stepfunctions is that one can use OLS, the disadvantage, is that parsimony may be gone.

Example I Revisited: Risk-Return Trade-off

- We estimate via QMLE the parameters θ_i jointly with μ and γ using MIDAS regression:

$$R_{t+1} \sim N\left(\mu + \gamma V_t^{\text{MIDAS}}, V_t^{\text{MIDAS}}\right) \tag{5}$$

where

$$V_t^{\text{MIDAS}} = 22 \sum_{d=1}^{\infty} w(d, \theta_1, \theta_2) r_{t-d}^2$$

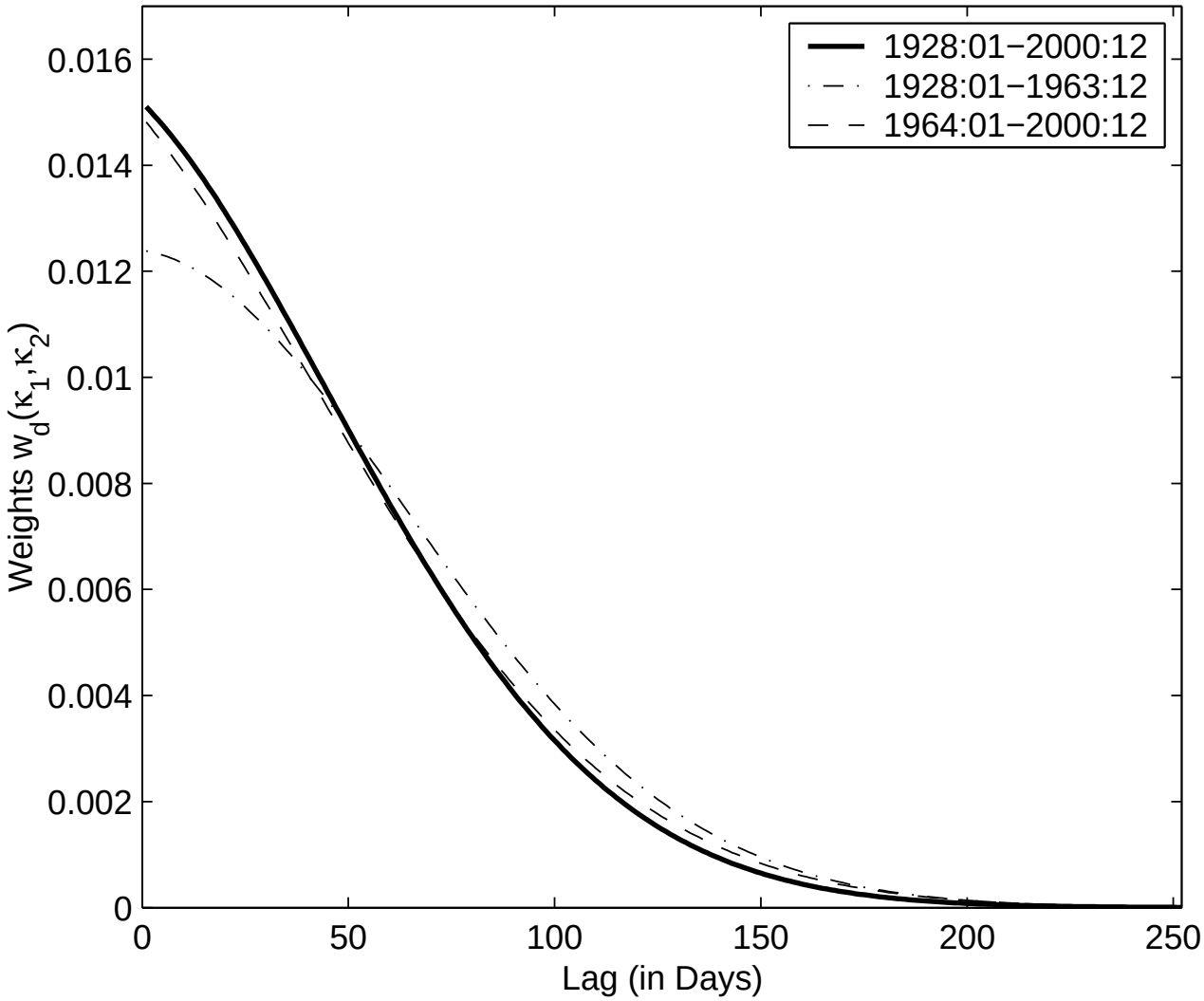
and

$$w(d, \theta_1, \theta_2) = \frac{\exp\{\theta_1 d + \theta_2 d^2\}}{\sum_{i=1}^{\infty} \exp\{\theta_1 i + \theta_2 i^2\}}$$

- In all the results that follow, we use the past 260 days as the maximum lag length (results are not sensitive to increasing the lag length beyond one year, nor to alternative polynomial specifications).

Risk Return Trade-off: MIDAS regression results

Sample	μ ($\times 10^3$)	γ	θ_1 ($\times 10^2$)	θ_2 ($\times 10^9$)	R^2_R	$R^2_{\sigma^2}$	LLF
1946.01-2000.12	4.800	4.007	-1.353	-3.984	0.024	0.082	1221.837
	[2.419]	[2.647]	[-1.903]	[-0.092]			
1946.01-2000.12	4.809	4.254	-1.402	-3.293	0.041	0.251	1239.100
(No 1987 Crash)	[2.515]	[2.950]	[-1.959]	[-0.011]			


1928:01−2000:12
1928:01−1963:12
1964:01−2000:12
Weights $w_d(\kappa_1,\kappa_2)$
Lag (in Days)
0
0.002
0.004
0.006
0.008
0.01
0.012
0.014
0.016
0
50
100
150
200
250

We estimate via QMLE the GARCH-M:

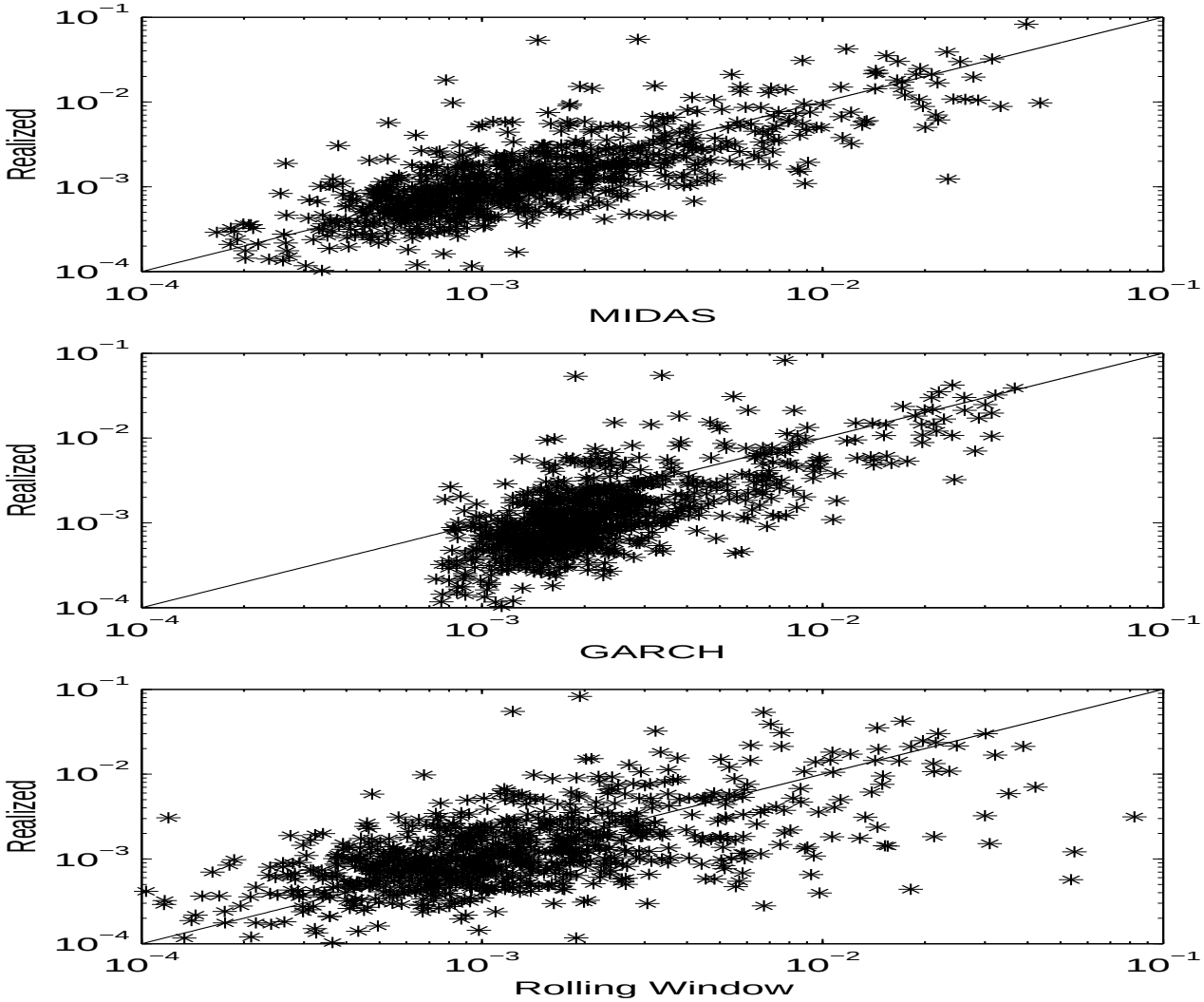
$$R_{t+1} \sim N\left(\mu + \gamma V_t^{\text{GARCH}}, V_t^{\text{GARCH}}\right)$$

where

$$V_t^{\text{GARCH}} = \frac{\omega}{1-\beta} + \alpha \sum_{i=0}^{\infty} \beta^i \epsilon_{t-i}^2$$

using past *monthly* squared returns. Results are consistent with Glosten, Jagannathan, and Runkle (1993) and others in the literature, namely:

Model	μ ($\times 10^3$)	γ	ω ($\times 10^3$)	α	β	R_R^2	$R_{\sigma^2}^2$	LLF
GARCH(1,1)-M	-0.740	6.968	0.125	0.069	0.860	0.010	0.070	1152.545
	[-0.370]	[0.901]	[0.244]	[1.398]	[18.323]			
ABS-GARCH(1,1)-M	1.727	6.013	2.751	0.099	0.858	0.010	0.071	1156.142
	[0.424]	[0.873]	[0.947]	[1.764]	[17.323]			


Realized
MIDAS
Realized
GARCH
Realized
Rolling Window

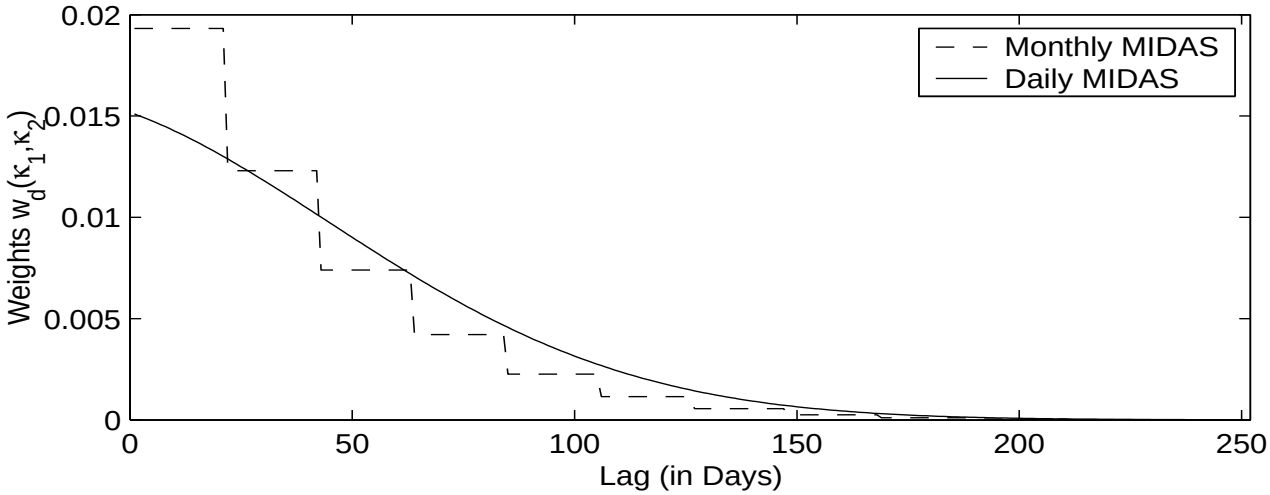
- The Kalman Filter can be used to interpolate "missing data" (see e.g. Harvey and Pierse (1984) and subsequent work). The Kalman Filter applies the linear Gaussian systems. In general settings, however, aggregation and interpolation is not so straightforward.
- Advantage of MIDAS is the reduced form approach.
- Along similar lines one can build up high-frequency data GARCH models and through temporal aggregation (see however, Drost and Nijman (1993) and Meddahi and Renault (2003)) to obtain MIDAS regression projection. But here too there are limitations, works for "simple" models.
- Suppose data is generated by two-factor GARCH model at $1/m$ frequency and one runs MIDAS regression:

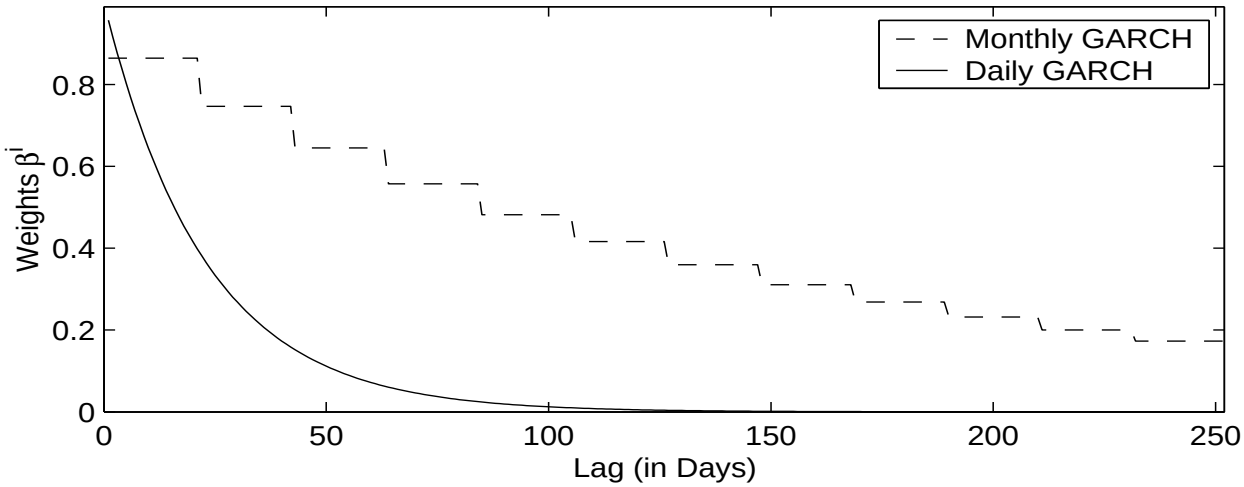
$$\sum_{j=1}^{m}[r_{t+j/m}^{(m)}]^2 = \beta_0 + \beta_1 B(L^{1/m})[r_t^{(m)}]^2 + \varepsilon_t$$

Reverse Engineering MIDAS projection two-factor volatility model

$$((\rho_{1(m)}+\rho_{2(m)})+(\rho_{1(m)}+\rho_{2(m)})^2+\rho_{1(m)}\rho_{2(m)}+(\rho_{1(m)}+\rho_{2(m)})^3+2(\rho_{1(m)}$$
$$+\rho_{2(m)})\rho_{1(m)}\rho_{2(m)}+(\rho_{1(m)}+\rho_{2(m)})^4+3(\rho_{1(m)}+\rho_{2(m)})^2\rho_{1(m)}\rho_{2(m)}$$
$$+(\rho_{1(m)}\rho_{2(m)})^2)+(\rho_{1(m)}\rho_{2(m)}+(\rho_{1(m)}+\rho_{2(m)})\rho_{1(m)}\rho_{2(m)}+(\rho_{1(m)}+\rho_{2(m)})^2\rho_{1(m)}\rho_{2(m)}$$
$$+(\rho_{1(m)}\rho_{2(m)})^2+(\rho_{1(m)}+\rho_{2(m)})^3\rho_{1(m)}\rho_{2(m)}+2(\rho_{1(m)}+\rho_{2(m)})(\rho_{1(m)}\rho_{2(m)})^2)L^{1/m}$$
$$+((\rho_{1(m)}\rho_{2(m)}-\rho_{1(m)}\alpha_{2(m)}-\rho_{2(m)}\alpha_{1(m)})-(\rho_{1(m)}+\rho_{2(m)})(\rho_{1(m)}+\rho_{2(m)}-\alpha_{1(m)}$$
$$-\alpha_{2(m)})-(\rho_{1(m)}+\rho_{2(m)}-\alpha_{1(m)}-\alpha_{2(m)})-(\rho_{1(m)}+\rho_{2(m)})^2(\rho_{1(m)}+\rho_{2(m)}-\alpha_{1(m)}$$
$$-\alpha_{2(m)})-\rho_{1(m)}\rho_{2(m)}(\rho_{1(m)}+\rho_{2(m)}-\alpha_{1(m)}-\alpha_{2(m)})+(\rho_{1(m)}+\rho_{2(m)})(\rho_{1(m)}\rho_{2(m)}$$
$$-\rho_{1(m)}\alpha_{2(m)}-\rho_{2(m)}\alpha_{1(m)})+(\rho_{1(m)}-\rho_{2(m)})^3(\rho_{1(m)}+\rho_{2(m)}$$
$$-\alpha_{1(m)}-\alpha_{2(m)})+(\rho_{1(m)}+\rho_{2(m)})^2(\rho_{1(m)}\rho_{2(m)}-\rho_{1(m)}\alpha_{2(m)}-\rho_{2(m)}\alpha_{1(m)})$$
$$-2(\rho_{1(m)}+\rho_{2(m)})\rho_{1(m)}\rho_{2(m)}(\rho_{1(m)}+\rho_{2(m)}-\alpha_{1(m)}-\alpha_{2(m)})$$
$$+\rho_{1(m)}\rho_{2(m)}(\rho_{1(m)}\rho_{2(m)}-\rho_{1(m)}\alpha_{2(m)}-\rho_{2(m)}\alpha_{1(m)}))(1-(\rho_{1(m)}+\rho_{2(m)})L^{1/m}$$
$$+\rho_{1(m)}\rho_{2(m)}L^{2/m})/(1-(\rho_{1(m)}+\rho_{2(m)}-\alpha_{1(m)}-\alpha_{2(m)})L^{1/m}+(\rho_{1(m)}\rho_{2(m)}$$
$$-\rho_{1(m)}\alpha_{2(m)}-\rho_{2(m)}\alpha_{1(m)})L^{2/m})+((\rho_{1(m)}\rho_{2(m)}-\rho_{1(m)}\alpha_{2(m)}-\rho_{2(m)}\alpha_{1(m)})$$
$$+\rho_{1(m)}\rho_{2(m)}(\rho_{1(m)}\rho_{2(m)}-\rho_{1(m)}\alpha_{2(m)}-\rho_{2(m)}\alpha_{1(m)})+$$
$$(\rho_{1(m)}+\rho_{2(m)})^2(\rho_{1(m)}\rho_{2(m)}-\rho_{1(m)}\alpha_{2(m)}-\rho_{2(m)}\alpha_{1(m)})+\rho_{1(m)}\rho_{2(m)}(\rho_{1(m)}\rho_{2(m)}$$
$$-\rho_{1(m)}\alpha_{2(m)}-\rho_{2(m)}\alpha_{1(m)})+(\rho_{1(m)}+\rho_{2(m)})^3$$
$$+2(\rho_{1(m)}+\rho_{2(m)})\rho_{1(m)}\rho_{2(m)}(\rho_{1(m)}\rho_{2(m)}-\rho_{1(m)}\alpha_{2(m)}-\rho_{2(m)}\alpha_{1(m)}))\times$$
$$(1-(\rho_{1(m)}+\rho_{2(m)})L^{1/m}+\rho_{1(m)}\rho_{2(m)}L^{2/m})L^{1/m}/$$
$$(1-(\rho_{1(m)}+\rho_{2(m)}-\alpha_{1(m)}-\alpha_{2(m)})L^{1/m}+(\rho_{1(m)}\rho_{2(m)}-\rho_{1(m)}\alpha_{2(m)}$$
$$-\rho_{2(m)}\alpha_{1(m)})L^{2/m})$$

Comparison of MIDAS and GARCH using Daily and Monthly Returns





Asymmetries

To examine whether the risk-return tradeoff is robust to the inclusion of asymmetric effects in the conditional variance, we introduce the asymmetric MIDAS estimator:

$$V_t^{\text{ASYMIDAS}} = 22[\phi \sum_{d=1}^{\infty} w(d, \theta_1^-, \theta_2^-)\mathbf{1}_{t-d}^- r_{t-d}^2 + (2-\phi) \sum_{d=1}^{\infty} w(d, \theta_1^+, \theta_2^+)\mathbf{1}_{t-d}^+ r_{t-d}^2]$$

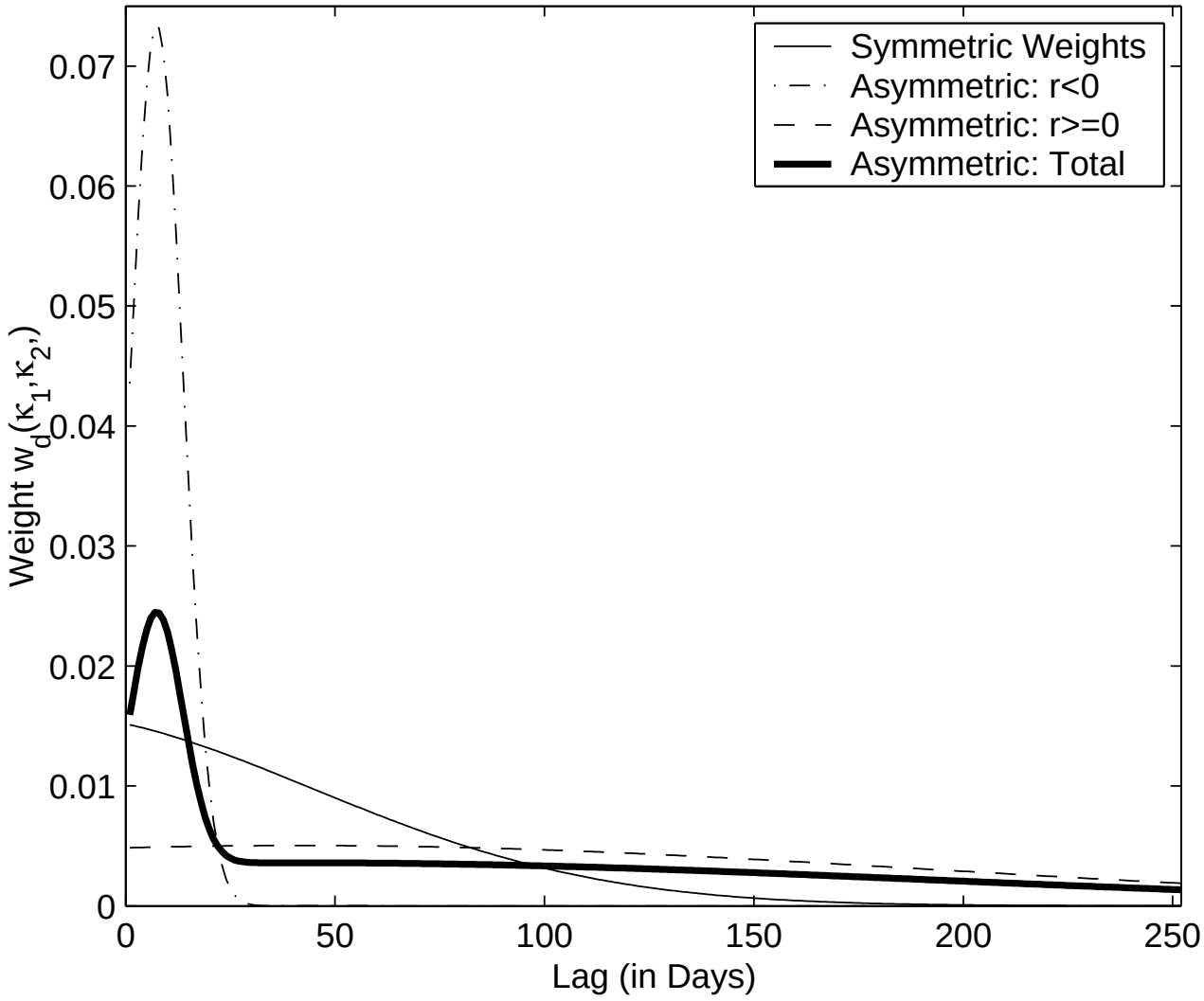
where $\mathbf{1}_{t-d}^+$ denotes the indicator function for $\{r_{t-d} \geq 0\}$, $\mathbf{1}_{t-d}^-$ denotes the indicator function for $\{r_{t-d} < 0\}$, and ϕ is in the interval $(0, 2)$.

Results with Asymmetries

Sample	μ ($\times 10^3$)	γ	+/-	θ_1 ($\times 10^2$)	θ_2 ($\times 10^3$)	ϕ	R^2_R	$R^2_{\sigma^2}$
1946:01-	5.766	3.314	(-)	9.573	-7.640	0.606	0.025	0.085
2000:12	[2.057]	[2.695]		[0.507]	[-0.929]	[3.381]		
			(+)	0.073	-0.210			
				[0.133]	[-0.539]			
1946:01-	5.550	3.735	(-)	-7.541	-2.340	0.716	0.043	0.363
2000:12	[1.989]	[2.868]		[-1.775]	[-0.342]	[5.010]		
(No Crash)			(+)	-0.214	-0.910			
				[-0.613]	[-0.498]			

GARCH-M Comparison - results are consistent with Glosten, Jagannathan, and Runkle (1993) and others in the literature.

Model	μ ($\times 10^3$)	γ	ω ($\times 10^2$)	α	β	λ ($\times 10^2$)	R^2_R	$R^2_{\sigma^2}$	LLF
EGARCH(1,1)-M	14.978	-2.521	-640.708	-0.325	0.497	-3.339	0.011	0.071	1159.102
	[6.277]	[-1.285]	[-1.790]	[-2.977]	[5.938]	[-2.206]			
ASYGARCH(1,1)-M	1.117	-3.248	0.056	0.018	0.609	-28.723	0.010	0.077	1164.023
	[0.913]	[-1.811]	[0.202]	[1.980]	[7.842]	[-2.131]			
QGARCH(1,1)-M	13.970	-1.994	0.060	0.086	0.145	-9.320	0.010	0.072	1161.173
	[2.378]	[-0.171]	[0.356]	[3.565]	[3.269]	[-7.188]			


Symmetric Weights
Asymmetric: r<0
Asymmetric: r>=0
Asymmetric: Total
Weight $w_d(\kappa_1,\kappa_2,)$
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
50
100
150
200
250
Lag (in Days)

Example II Revisited: Predicting Realized Volatility

- The objective of interest is predicting the increments in the quadratic variation of the return process over some future period, H, from one week ($H = 5$) to one month ($H = 20$) horizon. This is the variance that matters for option pricing and portfolio management.
- The quadratic variation is not observed directly but can be measured with some discretization error. One such measure would be the sum of (future) (intra-daily) squared returns.

- We start with $daily$ regressors and consider the following:

$$\tilde{Q}_{t+H,t}^{(Hm)} = \mu_H + \phi_H \sum_{k=0}^{k^{max}} b_H^Q(k,\theta) X_{t-k} + \varepsilon_{Ht}^X$$

  where the following regressors are considered:
  - Past daily quadratic variation $\tilde{Q}_{t,t-1}^{(m)}$ advocated by Andersen et al. (2001, 2002, 2003). The daily QV is computed by the sum of (say) 5-minute intra-daily squared returns.
  - Past daily squared returns, corresponds to the ARCH/GARCH class of models (under some parameter restrictions).
  - Past daily absolute returns.
  - Past daily high minus low (advantage is that it is available over long periods, unlike HF data.
  - Past $Power\ variation$, i.e. sum of (say) 5-minute intra-daily **absolute** returns, see in particular Barndorff-Nielsen and Shephard (2002c) and Woerner (2002).

- All of the MIDAS regressions come in various 'flavors', i.e. using log transformations, or square root etc.
- Andersen et al. (2003) advocate the use of long memory models to parsimoniously parameterize the weights. In particular, they consider models of the following type:

$$(1 - \sum_{k=1}^{5} b_A(k) L^k)(1 - L)^d \log \tilde{Q}_{t+1,t}^{(m)} = \mu + \varepsilon_t \qquad (6)$$

- This model will be our benchmark for all in-sample and out-of-sample forecast comparisons and is henceforth referred to as the "ABDL" model.
- Dataset consists of five-minute intra-day returns of the Dow Jones Composite Portfolio (DJ) over a ten year period, from April 1, 1993 to October 31, 2003.

In-Sample MSE Comparisons of MIDAS Models with Daily Regressors - DJ Index

(Ratio of MSE's w.r.t. ABDL Sample Apr. 1, 93 - Mar. 31, 01

	$\tilde{Q}^{(m)}_{t,t-1}$	$r^2_{t,t-1}$	$\lvert r_{t,t-1}\rvert$	$[hi-lo]_{t,t-1}$	$\tilde{P}^{(m)}_{t,t-1}$
Panel A: $\tilde{Q}^{(m)}_{t,t-H}$ MIDAS with daily lags of regressors					
1 wk	0.871	1.043	0.969	0.856	0.825
2 wks	0.851	0.962	0.875	0.791	0.758
3 wks	0.793	0.881	0.774	0.722	0.691
4 wks	0.805	0.908	0.806	0.743	0.681

	$\log(\tilde{Q}^{(m)}_{t,t-1})$	$\log(r_{t,t-1})^2$	$\log\lvert r_{t,t-1}\rvert$	$\log[hi-lo]_{t,t-1}$	$\log(\tilde{P}^{(m)}_{t,t-1})$
Panel B: $\log(\tilde{Q}^{(m)}_{t,t-H})$ MIDAS with daily lags of regressors					
1 wk	0.924	1.509	—	1.062	0.921
2 wks	0.881	1.396	—	1.015	0.867
3 wks	0.821	1.134	—	0.893	0.772
4 wks	0.812	1.177	—	0.921	0.761

Out-of-Sample MSE Comparisons of MIDAS Models with Daily Regressors - DJ Index

	$\tilde{Q}^{(m)}_{t,t-H}$ MIDAS					log($\tilde{Q}^{(m)}_{t,t-H}$) MIDAS			
	$\tilde{Q}^{(m)}_{t,t-1}$	$r^2_{t,t-1}$	$\lvert r_{t,t-1}\rvert$	$[hi-lo]_{t,t-1}$	$\tilde{P}^{(m)}_{t,t-1}$	log $\tilde{Q}^{(m)}_{t,t-1}$	log $r^2_{t,t-1}$	log$[hi-lo]_{t,t-1}$	log $\tilde{P}^{(m)}_{t,t-1}$
1 wk	0.802	1.048	1.005	0.819	0.778	0.746	1.457	0.815	0.729
2 wks	0.920	1.168	1.039	0.855	0.726	0.794	1.296	0.882	0.731
3 wks	0.814	0.995	0.874	0.794	0.724	0.751	1.089	0.797	0.714
4 wks	0.872	1.089	0.985	0.857	0.774	0.893	1.162	0.981	0.854

Some Theory

- A typical continuous time SV model for log-prices $p(t)$ can be written as:

$$dp\left(t\right)=\mu\left(t\right)dt+\sigma\left(t\right)dW\left(t\right) \tag{7}$$

- To make the comparison of persistence properties, it will be most convenient to specify a diffusion for $\sigma(t)$ and derive the implied autocorrelation properties of quadratic and power variation.
- Following Barndorff-Nielsen and Shephard (2001) we use a non-Gaussian Ornstein-Uhlenbeck (OU) process:

$$d\sigma(t)=-\lambda\sigma(t)dt+dz(\lambda t) \tag{8}$$

where $z(t)$ is a Lévy process with non-negative increments.

- Within this diffusion framework Forsberg and Ghysels (2004) show that PV *is expected to be* the best predictor. We expect: $PV\succ BPV\left(C\right)\succ QV.$

Models - MIDAS with step functions

We predict the normalized Realized Variance in-sample and out-of-sample using

- MIDAS regressions with polynomials (Ghysels, Santa-Clara and Valkanov(2003a JFE forthcoming, 2003b JoE forthcoming))

$$RV_{t,t+H}^{1/2} = \mu_H + B(L) X_t + \varepsilon_{Ht}$$

- MIDAS with stepfunctions Forsberg and Ghysels (2004) (special case is HAR (Heterogenous Autoregressive) Model (Corsi(2003)))

$$RV_{t,t+H}^{1/2} = \beta_0 + \beta_D X_{t-1,t} + \beta_W X_{t-5,t} + \beta_M X_{t-20,t} + \varepsilon_{t+H}$$

In-Sample Results modeling $RV^{1/2}$ of S&P 500 1985-2003 - Sign level of the Bipower test = 0.999

	Stepfunction MIDAS-RV$^{1/2}$					Beta polynomial MIDAS-RV$^{1/2}$				
Horizon	RV$^{1/2}$	BPV$^{1/2}$	C$^{1/2}$	$(CJ)^{1/2}$	RAV	RV$^{1/2}$	BPV$^{1/2}$	C$^{1/2}$	$(CJ)^{1/2}$	RAV
R2										
1 day	0.591	0.592	0.592	0.594	0.623	0.596	0.595	0.598	0.599	0.617
1 wk	0.656	0.649	0.654	0.658	0.690	0.671	0.667	0.671	0.672	0.687
2 wks	0.648	0.639	0.646	0.651	0.690	0.670	0.665	0.670	0.671	0.689
3 wks	0.633	0.623	0.631	0.636	0.680	0.656	0.651	0.656	0.657	0.678
4 wks	0.615	0.603	0.613	0.617	0.665	0.638	0.632	0.638	0.639	0.662
MSE										
1 day	1.344	1.333	1.343	1.339	1.265	1.314	1.314	1.311	1.309	1.269
1 wk	0.513	0.519	0.517	0.510	0.458	0.482	0.486	0.481	0.480	0.463
2 wks	0.382	0.391	0.385	0.378	0.330	0.350	0.355	0.349	0.348	0.331
3 wks	0.339	0.349	0.342	0.335	0.290	0.310	0.316	0.309	0.308	0.292
4 wks	0.319	0.330	0.323	0.316	0.274	0.293	0.300	0.294	0.292	0.277

Summary and robustness of empirical results

- RAV is the best predictor/ regressor for RV (as predicted by theory)
- Robustness of empirical results:
  - Result holds when modeling Realized Variance in levels
  - Result is the same regardless of the significance level of the bipower test, we have investigated $\alpha = 0.5$, 0.95, 0.99 and 0.999
  - Same results when a dummy for the jump-days is included
  - Robust to subsample period: S&P 1990-2002
  - Results are robust to using other evaluation measures such as Heterscedasticity Adjusted Error (Bollerslev and Ghysels (1996))

Example III Revisited: Impact of news

Work in progress joint with Arthur Sinko and Ross Valkanov

- The realized return of an asset between periods $t-1$ and t can always be written as

$$r_{i,t} = \mathcal{E}_{t-1}(r_{i,t}) + \varepsilon_{i,t} \tag{9}$$

  where asset i's expected return $\mathcal{E}_{t-1}(r_{i,t})$ is obtained using all available information at time $t-1$.

- The second term, $\varepsilon_{i,t}$, is the unexpected return which reflects the response of market participants to news in that period.
- Ghysels, Sinko and Valkanov (2005) document the impact of macro and corporate news on the entire cross section of individual stock returns.
- They use a reverse MIDAS regressions which allow them to analyze the lead-lag relation between variables sampled at different frequencies.

- Recall Granger causality, meaning the one-step ahead prediction of y_t using past y_{t-i} can be improved upon in a least squares sense by using also lagged x_{t-i}.
- While there are many ways of conducting Granger causality tests, Sims proposed the following regression:

$$x_t = a_0 + \sum_{j=1}^{k_F} b_j^F y_{t+j} + \sum_{j=1}^{k_P} b_j^P y_{t-j} + \varepsilon_t$$

- Testing for Granger causality between x_t and y_t amounts to testing whether the coefficients b_j^F are equal to zero.
- With this idea in mind let us revisit MIDAS regressions and let us consider:

$$x_t = \beta_0 + \beta_1^F B_F(L^{-1/m}) y_t^{(m)} + \beta_1^P B_P(L^{1/m}) y_t^{(m)} + \varepsilon_t \qquad (10)$$

where $B_F(L^{1/m}) = b_0^F + b_1^F L^{1/m} + b_2^F L^{2/m} + \ldots + b_{j^{max}}^F L^{j_F^{max}/m}$ is a polynomial of length j_F^{max} in the $L^{1/m}$ operator, and $L^{j/m} x_t^{(m)} = x_{t-j/m}^{(m)}$.

- The reverse MIDAS specification allows us to separate how much markets anticipated the announcement, via the parameter β_1^P and the polynomial $b_1^P L^{1/m}$ and how much markets responded to the announcement, via the parameter β_1^F and the polynomial $b_1^P F^{1/m}$.
- We look at daily and intra-daily projections.
- We examine the $entire$ cross-section of stocks.

- We find that macro news (GDP growth, industrial production, earnings per capita, consumer confidence, unemployment claims, CPI inflation, PPI inflation, and housing starts) have an impact on the cross section of stock returns. This is contrary to what previous studies have found (e.g., Patelis (1999)).
- Moreover, the impact of macro news differs across firm characteristics. For instance, returns of small market capitalization stocks respond more to macro news than the returns of medium or large cap stocks.
- In addition, returns of large book-to-market firms respond more to macro news.

- The economic and statistical significance of our findings is due to at least two factors:
  - First, reverse MIDAS regressions represent a statistically powerful way of capturing the impact of news on stock returns.
  - Second, by considering the entire cross section of publicly traded U.S. companies, we are able to identify the impact of macroeconomic news where previous studies focusing on limited number of firms might have failed.
- We also find that microeconomic (earnings) announcements have little effect on the cross section of returns.

Example IV Revisited: Application of MIDAS to forecasting the predictions of forecasters

Work in progress joint with Jonathan Wright

- MIDAS methodology ideally suited to using high frequency financial data to predict low frequency macro data
- Low frequency macro forecasts are less noisy than $ex\text{-}post$ realized macro data and so potentially easier to predict
- Once a month/quarter, observe forecasts of some future macro/finance variable (e.g. inflation four quarters hence), f_t, from
  - Survey of Professional Forecasters (quarterly)
  - Consensus Forecasts (monthly)
- Know deadline dates for completion of survey: d_t

- Actual timing of responses is however fuzzy
- Objective: use daily asset return data to predict the next forecast.
- Allows for forecasts to be predicted at daily frequency
- Useful to policy makers and analysts to know best guess for market forecast of next years' inflation as of today.
- Conceivably could even use intradaily data

- Model: Prediction equation day $d^*_{t+1} = \theta(d_{t+1} - d_t)$, $0 < \theta \leq 1$:

$$f_{t+1} - f_t = \alpha + \Sigma_{j=1}^{k} \beta_j \gamma(L) r_{t,j}^{DAILY} + \rho f_t + \varepsilon_t$$

  where $\gamma(L) r_{t,i}^{DAILY}$ is dist. lag of daily returns on asset i over n_l days up to d^*_{t+1}.

- Model can then be used to forecast f_{t+1} as of day d^*_{t+1} e.g. if $\theta = 1$, predicts f_{t+1} as of completion deadline date (survey not out yet)

- To avoid overparameterizing $\gamma(L)$, can use a MIDAS specification and we estimate parameters of the model $(\alpha, \beta_1, \beta_2, ...\beta_k, \rho, \kappa_1, \kappa_2)$ by MLE

- Can compare three predictions:
  - Prediction from the baseline model with no asset returns at all:

$$f_{t+1} - f_t = \alpha + \rho f_t + \varepsilon_t$$

    estimated by OLS (model BASELINE).
  - Prediction from the model using the average daily asset returns over n_l days prior to d^*_{t+1} but without estimating any MIDAS polynomial, imposing : $\kappa_1 = \kappa_2 = 1$
  - or equivalently $\gamma(L) = \sum_{j=1}^{n_l} n_l^{-1} L^{j-1}$ (model EW-MIDAS).
  - And finally, prediction from the full-blown MIDAS model (MIDAS).
- For asset returns use, for example,
  - excess stock market returns and
  - changes in the fourth eurodollar futures contract.

Compare the R-squareds from the three prediction methods

Prediction		$\theta = 1$		$\theta = 2/3$		$\theta = 1/3$	
	BASELINE	EW-MIDAS	MIDAS	EW-MIDAS	MIDAS	EW-MIDAS	MIDAS
SPF ($n_l = 90$)							
1 Quarter Hence							
GDP	0.20	0.47	0.58	0.49	0.59	0.51	0.58
CPI	0.15	0.29	0.34	0.20	0.26	0.21	0.28
T Bill	0.05	0.60	0.68	0.65	0.68	0.59	0.64
Unemployment	0.02	0.37	0.44	0.41	0.45	0.43	0.48
Profits	0.14	0.28	0.28	0.19	0.33	0.20	0.31
4 Quarters Hence							
GDP	0.07	0.13	0.21	0.11	0.24	0.11	0.25
CPI	0.05	0.24	0.36	0.60	0.68	0.65	0.68
T Bill	0.05	0.58	0.70	0.66	0.71	0.56	0.64
Unemployment	0.02	0.39	0.53	0.44	0.54	0.45	0.57
Profits	0.24	0.32	0.32	0.29	0.32	0.27	0.32

Asymptotics

I: Aggregation Bias and Aliasing Revisited

- When data of different sampling frequencies are mixed, one invariably deals with temporal aggregation.
- To study aggregation issues it is convenient to assume that the underlying stochastic processes evolve in continuous time and data are collected at discrete points in time.
- Such a setting has the appeal of imposing a priori a structure on discretely observed data that is independent of the sampling interval.

- To discuss many issues ranging from parameterization and approximations to discretization biases let us start with the continuous time setting:

$$\begin{aligned} y(t) &= b * x(t) + u(t) \\ &= \int_{-\infty}^{\infty} x(t-s)b(s)ds + u(t) \end{aligned} \tag{11}$$

  where the symbol $*$ denotes the convolution operator. The errors in equation (**??**) are not necessarily i.i.d.

- Identification of b in equation (**??**) rests on the assumption that the x process is, up to second moments, truly exogenous, i.e. $E[x(t)u(s)] = 0$, $\forall$ s and t.

- Sims (1971) and Geweke (1978) examine equations like (**??**) and study the relationship between inference drawn from discrete time models and the parameters of the continuous time convolution.

- A discrete time *distributed lag* model corresponding to (**??**) would be as follows:

$$Y_{t/m}^{(m)} = \frac{1}{m} \sum_{s=-\infty}^{\infty} B^{(m)}(\frac{s}{m}) X_{(t-s)/m}^{(m)} + U_{t/m}^{(m)} \tag{12}$$

  where both y and x are sampled at frequency $1/m$.

- The topic of discretization bias in distributed lag models, i.e. the difference between an estimator $B^{(m)}$ and b for any given m, has been extensively studied, see for instance Sims (1971), Geweke (1978), Hansen and Sargent (1983, 1991), Hansen and Sargent (1991), Phillips (1972, 1973, 1974), among others.

- MIDAS regressions involve processes with various sampling frequencies. More specifically, we study projections of Y sampled with $m = 1$ and $X^{(m)}$ sampled with $m > 1$. MIDAS regression models are therefore:

$$Y_t = \frac{1}{m} \sum_{s=-\infty}^{\infty} \bar{B}^{(m)}(\frac{s}{m}) X^{(m)}_{(t-s)/m} + U_t \tag{13}$$

- It is important to note that we only deal with OLS estimators, and therefore are not interested at this stage with efficiency issues. Hence, we examine OLS estimators $B^{(m)}$ in distributed lag models, similar to Sims (1971) and Geweke (1978), and OLS estimators $\bar{B}^{(m)}$ in MIDAS regressions.

- The discretely sampled distributed lag regression yields the OLS estimator:

$$\tilde{B}^{(m)} = F_m[S_x\tilde{b}]/F_m[S_x] = F_m[S_{yx}]/F_m[S_x] \tag{14}$$

  where S_{yx} is the co-spectrum of continuously sampled $y(t)$ and $x(t)$.

- The intuition why equation (**??**) also suggests that MIDAS regressions may resemble distributed lag models in terms of discretization bias, it is important to note that what matters, besides $F_m[S_x]$, is the covariance $F_m[S_{yx}]$. In a MIDAS regression, assuming stationarity and point sampling of y and x it is clear that ultimately we recover the covariance between y_t and any lag of x_t.

II: Efficiency comparisons between MIDAS and distributed lag models

- Consider again the discrete time $distributed\ lag$ model like (**??**) where both y and x are sampled at a $fixed$ frequency $1/m$:

$$Y_{t/m}^{(m)} = \frac{1}{m} \sum_{s=-\infty}^{\infty} b^{(m)}(\frac{s}{m}) X_{(t-s)/m}^{(m)} + u_{t/m}^{(m)} \tag{15}$$

  where $b^{(m)}$ is the pseudo-true value associated with the fixed m. We try to obtain an $efficient$ estimator which we will denote B_H^m given a data set of size mT for both $Y^{(m)}$) and $X^{(m)}$.

- Consider also the MIDAS regression:

$$Y_t = \frac{1}{m} \sum_{s=-\infty}^{\infty} \overline{b}^{(m)}(\frac{s}{m}) X_{(t-s)/m}^{(m)} + u_t \tag{16}$$

- A simple strategy that leads to efficient estimation is to prefilter the equation by b_2 :

$$Y_{t/m}^{(m)} = \sum_{s=-\infty}^{\infty} (\tilde{b}_2^{(m)}(\frac{s}{m}))Y_{(t-s-1)/m}^{(m)} + \sum_{s=-\infty}^{\infty} b_1^{(m)}(\frac{s}{m})X_{(t-s)/m}^{(m)} + v_{t/m}^{(m)}$$

  where the availability of lagged $Y_{t/m}^{(m)}$ allows us to apply the polynomial b_2.

- In a MIDAS regression this strategy is infeasible due to the lack of high frequency $Y_{t/m}^{(m)}$. Consequently, the errors remain correlated and the estimator has to settle with an autocorrelation structure that cannot be further unraveled. The clear advantage of distributed lag models is the availability of the additional information about $Y^{(m)}$.