B Web supplement to Pesaran, Pick, Pranovich (2012)

This web supplement provides details that were omitted from the paper for brevity.

B.1 Weights for continuous break for extreme values of δ

Here, we derive the weights in (3) for $\delta = 0$ and $\delta \to \infty$. First, for $\delta = 0$ we have that $\delta^2 \mathbf{H}\mathbf{H}' + \mathbf{I}_T = \mathbf{I}_T$ and therefore

$$\theta = \frac{1}{\boldsymbol{\iota}_T'\boldsymbol{\iota}_T} = \frac{1}{T}.$$

The weights are therefore

$$\mathbf{w} = \frac{1}{T}\boldsymbol{\iota}_T.$$

Second, for $\delta \to \infty$, first rewrite the weights as

$$\mathbf{w} = (\mathbf{H}\mathbf{H}' + \delta^{-2}\mathbf{I})^{-1}\left(\mathbf{H} + \delta^{-2}\theta\mathbf{I}\right)\boldsymbol{\iota}_T,$$

and note that

$$\delta^{-2}\theta = \frac{\left[1 - \boldsymbol{\iota}_T'(\mathbf{H}\mathbf{H}' + \delta^{-2}\mathbf{I})^{-1}\mathbf{H}\boldsymbol{\iota}_T\right]}{\boldsymbol{\iota}_T'(\mathbf{H}\mathbf{H}' + \delta^{-2}\mathbf{I})^{-1}\boldsymbol{\iota}_T}.$$

In the case where δ is sufficiently large we can expand $(\mathbf{H}\mathbf{H}' + \delta^{-2}\mathbf{I})^{-1}$ as

$$\begin{aligned}
(\mathbf{H}\mathbf{H}' + \delta^{-2}\mathbf{I})^{-1} &= \mathbf{H}'^{-1}\left[\mathbf{I} + \delta^{-2}\left(\mathbf{H}'\mathbf{H}\right)^{-1}\right]^{-1}\mathbf{H}^{-1} \\
&= \mathbf{H}'^{-1}\left[\mathbf{I} - \delta^{-2}\left(\mathbf{H}'\mathbf{H}\right)^{-1} + \delta^{-4}\left(\mathbf{H}'\mathbf{H}\right)^{-2} - \delta^{-6}\left(\mathbf{H}'\mathbf{H}\right)^{-3} + \ldots\right]\mathbf{H}^{-1} \\
&= \left(\mathbf{H}\mathbf{H}'\right)^{-1} - \delta^{-2}\mathbf{H}'^{-1}\left(\mathbf{H}'\mathbf{H}\right)^{-1}\mathbf{H}^{-1} + \delta^{-4}\mathbf{H}'^{-1}\left(\mathbf{H}'\mathbf{H}\right)^{-2}\mathbf{H}^{-1} + \ldots
\end{aligned}$$

Hence

$$1 - \boldsymbol{\iota}_T'(\mathbf{H}\mathbf{H}' + \delta^{-2}\mathbf{I})^{-1}\mathbf{H}\boldsymbol{\iota}_T = 1 - \boldsymbol{\iota}_T'\left(\mathbf{H}\mathbf{H}'\right)^{-1}\mathbf{H}\boldsymbol{\iota}_T + \delta^{-2}\boldsymbol{\iota}_T'\mathbf{H}'^{-1}\left(\mathbf{H}'\mathbf{H}\right)^{-1}\mathbf{H}^{-1}\mathbf{H}\boldsymbol{\iota}_T - \ldots$$

and

$$\delta^{-2}\theta = \frac{1 - \boldsymbol{\iota}_T'\left(\mathbf{H}\mathbf{H}'\right)^{-1}\mathbf{H}\boldsymbol{\iota}_T + o(\delta^{-2})}{\boldsymbol{\iota}_T'(\mathbf{H}\mathbf{H}' + \delta^{-2}\mathbf{I})^{-1}\boldsymbol{\iota}_T}.$$

Therefore, as $\delta^2 \to \infty$

$$\begin{aligned}
\delta^{-2}\theta &\to \frac{1 - \boldsymbol{\iota}_T'\left(\mathbf{H}\mathbf{H}'\right)^{-1}\mathbf{H}\boldsymbol{\iota}_T}{\boldsymbol{\iota}_T'(\mathbf{H}\mathbf{H}')^{-1}\boldsymbol{\iota}_T}, \\
\lim_{\delta^2 \to \infty}\mathbf{w}(\delta) &= (\mathbf{H}\mathbf{H}')^{-1}\left[\mathbf{H} + \left(\frac{\mathbf{1} - \boldsymbol{\iota}_\mathbf{T}'\left(\mathbf{H}\mathbf{H}'\right)^{-\mathbf{1}}\mathbf{H}\boldsymbol{\iota}_\mathbf{T}}{\boldsymbol{\iota}_\mathbf{T}'(\mathbf{H}\mathbf{H}')^{-\mathbf{1}}\boldsymbol{\iota}_\mathbf{T}}\right)\mathbf{I}\right]\boldsymbol{\iota}_T.
\end{aligned}$$

But, using the fact that the triangular structure of $\mathbf{H}$ implies that

$$\mathbf{H}\mathbf{H}' = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & & 2 \\ 1 & 2 & 3 & & 3 \\ \vdots & & & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & T \end{pmatrix}, \text{ and } \left(\mathbf{H}\mathbf{H}'\right)^{-1} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & & 0 \\ 0 & -1 & 2 & & \vdots \\ \vdots & & & \ddots & -1 \\ 0 & 0 & & \cdots & 1 \end{pmatrix},$$

(see Neudecker, Trenkler and Liu, 2009, and Chu, Puntanem and Styan, 2011). Therefore,

$$\left(\mathbf{HH}'\right)^{-1}\mathbf{H}=\begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & & \vdots \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & & & \ddots & -1 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix},$$

$\left(\mathbf{HH}'\right)^{-1}\mathbf{H}\boldsymbol{\iota}_T=(0,\ldots,0,1)'$ and $\boldsymbol{\iota}_T'(\mathbf{HH}')^{-1}\mathbf{H}\boldsymbol{\iota}_T=1$. Also, $\boldsymbol{\iota}_T'(\mathbf{HH}')^{-1}\boldsymbol{\iota}_T=1$.

$$\lim_{\delta^2\to\infty}\mathbf{w}(\delta)=(\mathbf{HH}')^{-1}\mathbf{H}\boldsymbol{\iota}_T=(0,\ldots,0,1)'.$$

B.2 MSFE of post-break and optimal window

For the window that contains T_v of T observations the one-step ahead forecast is

$$\begin{aligned}\widehat{y}_{T+1} &= \frac{1}{T-T_v+1}\sum_{s=T_v}^{T}y_s = \frac{1}{T-T_v+1}\sum_{s=T_v}^{T_b}y_s+\frac{1}{T-T_v+1}\sum_{s=T_b+1}^{T}y_s \\ &= \frac{(T_b-T_v+1)\mu_{(1)}+\mu_{(2)}(T-T_b)}{T-T_v+1}+\frac{1}{T-T_v+1}\sum_{s=T_v}^{T}\sigma_\varepsilon\varepsilon_s.\end{aligned}$$

Set $v=\frac{T-T_v+1}{T}$ so that $T_v=T(1-v)+1$, and re-write the above as

$$\begin{aligned}\widehat{y}_{T+1}&=\mu_{(2)}\{1-\mathrm{I}[v-(1-b)]\}\\ &\quad+\mathrm{I}[v-(1-b)]\left\{\frac{(1-b)\mu_{(2)}+[v-(1-b)]\mu_{(1)}}{v}\right\}+\frac{1}{Tv}\sum_{s=T_v}^{T}\sigma_\varepsilon\varepsilon_s,\end{aligned}$$

where $\mathrm{I}(c)$ is an indicator function equal to 1 if $c>0$ and equal to 0 otherwise. The one-step ahead forecast error is

$$\widehat{e}_{T+1}=\left(\mu_{(2)}-\mu_{(1)}\right)\left[1-\frac{(1-b)}{v}\right]\mathrm{I}[v-(1-b)]+\sigma_\varepsilon\varepsilon_{T+1}-\frac{1}{Tv}\sum_{s=T_v}^{T}\sigma_\varepsilon\varepsilon_s.$$

The expected squared forecast error normalized by σ_ε^2 is

$$\begin{aligned}\mathrm{E}\left(\sigma_\varepsilon^{-2}\widehat{e}_{T+1}^2\right) &= 1+\frac{\left(\mu_{(2)}-\mu_{(1)}\right)^2}{\sigma^2}\left[1-\frac{(1-b)}{v}\right]^2\mathrm{I}[v-(1-b)]+\frac{1}{Tv}, \\ &= 1+\lambda^2\left[1-\frac{(1-b)}{v}\right]^2\mathrm{I}[v-(1-b)]+\frac{1}{Tv}. \qquad (56)\end{aligned}$$

Initially consider windows that do not contain the break. The window with all observations after the break will minimize the MSFE, so $v^o_{v\le(1-b)}=(1-b)$ and

$$\mathrm{E}\left[\sigma_\varepsilon^{-2}\widehat{e}_{T+1}^2|v=(1-b)\right]=1+\frac{1}{T(1-b)} \qquad (57)$$

This is also the MSFE of the forecast using the post-break window observations.

Now consider windows that include the break so that $\mathrm{I}[v-(1-b)]=1$ in (56). The first order condition is

$$\lambda^2\left[\frac{2(1-b)}{v^2}-\frac{2(1-b)^2}{v^3}\right]-\frac{1}{Tv^2}=0, \tag{58}$$

Then from (58), the expression for the optimal window (among those containing a break) is

$$v^o \;=\; \frac{2(1-b)^2\lambda^2}{2(1-b)\lambda^2-\frac{1}{T}} \;=\; (1-b)\frac{1}{1-\frac{1}{2\lambda^2(1-b)T}} \tag{59}$$

It can be seen that the optimal window is the distance to the break scaled by an expression that is larger the smaller the break and the smaller the distance to break. A condition of the optimal window is that it cannot exceed 1. Therefore $v^o=\frac{2(1-b)^2\lambda^2}{2(1-b)\lambda^2-\frac{1}{T}}\leqslant 1$, if $\lambda^2<\frac{T}{2(T-T_b)T_b}$ the optimal window contains all observations.

Using (59) in the MSFE (56) yields the results in (15).

B.3 Notes on the numerical solution of optimal weights in a multiple regression model

To this end let

$$\mathbf{d}(\mathbf{w})=\mathbf{S}^{-1}(\mathbf{w})\left(\sum_{t=1}^{T_b}q^2w_t^2\mathbf{x}_t\mathbf{x}_t'+\sum_{t=T_b+1}^{T}w_t^2\mathbf{x}_t\mathbf{x}_t'\right)\mathbf{S}^{-1}(\mathbf{w})\mathbf{x}_{T+1},$$

$$\mathbf{p}_1(\mathbf{w})=\mathbf{S}^{-1}(\mathbf{w})\boldsymbol{S}_1(\mathbf{w}_{(1)})\boldsymbol{\lambda},\quad \text{and} \quad \mathbf{p}_2(\mathbf{w})=\mathbf{S}^{-1}(\mathbf{w})\boldsymbol{S}_2(\mathbf{w}_{(2)})\boldsymbol{\lambda},$$

$$\gamma(\mathbf{w})=\mathbf{x}_{T+1}'\mathbf{S}^{-1}(\mathbf{w})\mathbf{S}_1(\mathbf{w}_{(1)})\boldsymbol{\lambda},$$

$$\theta_t(\mathbf{w})=\begin{cases} q^2\mathbf{x}_{T+1}'\mathbf{S}^{-1}(\mathbf{w})\mathbf{x}_t & \text{if } t\le T_b \\ \mathbf{x}_{T+1}'\mathbf{S}^{-1}(\mathbf{w})\mathbf{x}_t & \text{if } t>T_b, \end{cases},$$

where $\mathbf{d}(\mathbf{w})$, $\mathbf{p}_1(\mathbf{w})$ and $\mathbf{p}_2(\mathbf{w})$ are $k\times 1$ vectors, and $\gamma(\mathbf{w})$ and $\theta_t(\mathbf{w})$ are scalar functions of $\mathbf{w}$. Then, the T equations in (21) and (22) can be written as

$$\begin{aligned}\boldsymbol{\theta}(\mathbf{w})\odot\mathbf{w} &= \mathbf{X}\mathbf{d}(\mathbf{w})+\gamma(\mathbf{w})\begin{pmatrix}-\mathbf{X}_{(1)}\mathbf{p}_2(\mathbf{w})\\ \mathbf{X}_{(2)}\mathbf{p}_1(\mathbf{w})\end{pmatrix}\\ &= \mathbf{X}\mathbf{d}(\mathbf{w})+\gamma(\mathbf{w})\mathbf{Z}(\mathbf{w}),\end{aligned}$$

where $\mathbf{X}$ is the $T\times k$ matrix of regressors, $\mathbf{X}=\left(\mathbf{X}_{(1)}',\mathbf{X}_{(2)}'\right)'$, $\mathbf{X}_{(1)}$ and $\mathbf{X}_{(2)}$ are $T_b\times k$ and $(T-T_b)\times k$ matrices of pre-break and post-break regressors. Also, $\boldsymbol{\theta}(\mathbf{w})=(\theta_1(\mathbf{w}),\theta_2(\mathbf{w}),\ldots,\theta_T(\mathbf{w}))'$ and $\odot$ denotes element by element vector multiplication. We now need to minimize the function

$$\min_{\mathbf{w}}\mathbf{f}'(\mathbf{w})\mathbf{f}(\mathbf{w}),$$

subject to $\boldsymbol{\iota}_T'\mathbf{w}=1$ and $w_t\geq 0$, where

$$\mathbf{f}(\mathbf{w})=\boldsymbol{\theta}(\mathbf{w})\odot\mathbf{w}-\mathbf{X}\mathbf{d}(\mathbf{w})-\gamma(\mathbf{w})\mathbf{Z}(\mathbf{w}).$$

The asymptotic weights given in Section 2.3.2 below can be used as starting values for the numerical optimization.

B.4 Simple computation of weights for multiple breaks

Here we derive a simple representation of the optimal weights in matrix notation that is easy to implement in matrix oriented programming languages for any number of breaks in the mean and the variance.

Using the case discussed in Section 2.4.2, we have the first order conditions

$$
\begin{aligned}
t \leq T_1: \ & \phi_{(1)}[\phi_{(1)}\sum_{t=1}^{T_1} w_t + \phi_{(2)}\sum_{t=T_1+1}^{T_2} w_t + \cdots + \phi_{(n)}\sum_{t=T_{n-1}+1}^{T_n} w_t] + w_t q_t^2 + \theta/2 = 0, \\
T_1 < t \leq T_2: \ & \phi_{(2)}[\phi_{(1)}\sum_{t=1}^{T_1} w_t + \phi_{(2)}\sum_{t=T_1+1}^{T_2} w_t + \cdots + \phi_{(n)}\sum_{t=T_{n-1}+1}^{T_n} w_t] + w_t q_t^2 + \theta/2 = 0, \\
& \vdots \\
T_{n-1} < t \leq T_n: \ & \phi_{(n)}[\phi_{(1)}\sum_{t=1}^{T_1} w_t + \phi_{(2)}\sum_{t=T_1+1}^{T_2} w_t + \cdots + \phi_{(n)}\sum_{t=T_{n-1}+1}^{T_n} w_t] + w_t q_t^2 + \theta/2 = 0.
\end{aligned}
$$

We can write this in matrix notation as

$$(\mathbf{Q} + \boldsymbol{\phi}\boldsymbol{\phi}')\mathbf{w} = -\frac{\theta}{2}\boldsymbol{\iota},$$

where $\mathbf{Q}$ is a diagonal matrix with q_t^2 as the t,t-element, $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_T)'$, and

$$
\phi_t = \begin{cases}
\phi_{(1)} = \frac{\mathbf{x}_{T+1}\boldsymbol{\lambda}_1}{\mathbf{x}_{T+1}\boldsymbol{\Omega}_{xx}^{-1}\mathbf{x}_{T+1}} & \text{for } 1 \leq t \leq T_1 \\
\phi_{(2)} = \frac{\mathbf{x}_{T+1}\boldsymbol{\lambda}_2}{\mathbf{x}_{T+1}\boldsymbol{\Omega}_{xx}^{-1}\mathbf{x}_{T+1}} & \text{for } T_1 < t \leq T_2 \\
\qquad \vdots & \\
\phi_{(n)} = \frac{\mathbf{x}_{T+1}\boldsymbol{\lambda}_n}{\mathbf{x}_{T+1}\boldsymbol{\Omega}_{xx}^{-1}\mathbf{x}_{T+1}} & \text{for } T_{n-1} < t \leq T_n \\
\qquad 0 & T_n < t \leq T+1
\end{cases}
$$

Setting $\mathbf{M} = \mathbf{Q} + \boldsymbol{\phi}\boldsymbol{\phi}'$ we have

$$\mathbf{M}\mathbf{w} = -\frac{\theta}{2}\boldsymbol{\iota},$$

and therefore

$$\mathbf{w} = -\frac{\theta}{2}\mathbf{M}^{-1}\boldsymbol{\iota}. \tag{60}$$

Summing the weights over t yields

$$\boldsymbol{\iota}'\mathbf{w} = 1 = -\frac{\theta}{2}\boldsymbol{\iota}'\mathbf{M}^{-1}\boldsymbol{\iota}.$$

Hence,

$$-\frac{\theta}{2} = \frac{1}{\boldsymbol{\iota}'\mathbf{M}^{-1}\boldsymbol{\iota}}, \tag{61}$$

Using (61) in (60) yields the optimal weights

$$\mathbf{w} = \frac{1}{\boldsymbol{\iota}_T'\mathbf{M}^{-1}\boldsymbol{\iota}_T}\mathbf{M}^{-1}\boldsymbol{\iota}_T. \tag{62}$$

B.5 MSFE for robust optimal weights

Consider the MSFE associated with the robust optimal weights defined in (44). For these weights we need to compute $\sum_{t=1}^{T_b} w_t$, $\sum_{t=1}^{T_b} w_t^2$, and $\sum_{t=1}^{T} w_t^2$. Note that when T and T_b are relatively large we can use the following approximations (noting that by assumption $\underline{b} \leq b \leq \bar{b}$)

$$\sum_{t=1}^{T_b} w_t \approx \frac{-1}{\left(\bar{b}-\underline{b}\right)} \int_{\underline{b}}^{b} \log\left(\frac{1-a}{1-\underline{b}}\right) da,$$

$$\sum_{t=1}^{T_b} w_t^2 \approx \frac{1}{T\left(\bar{b}-\underline{b}\right)^2} \int_{\underline{b}}^{b} \left[\log\left(\frac{1-a}{1-\underline{b}}\right)\right]^2 da,$$

$$\sum_{t=1}^{T} w_t^2 \approx \frac{1}{T\left(\bar{b}-\underline{b}\right)^2} \int_{\underline{b}}^{\bar{b}} \left[\log\left(\frac{1-a}{1-\underline{b}}\right)\right]^2 da + \frac{(1-\bar{b})}{T\left(\bar{b}-\underline{b}\right)^2} \left[\log\left(\frac{1-\bar{b}}{1-\underline{b}}\right)\right]^2.$$

First, note that

$$\sum_{t=1}^{T_b} w_t \approx \frac{-1}{\left(\bar{b}-\underline{b}\right)} \int_{\underline{b}}^{b} \log(1-a)da + \frac{(b-\underline{b})}{\left(\bar{b}-\underline{b}\right)} \log(1-\underline{b}) = \frac{b-\underline{b}}{\bar{b}-\underline{b}} + \frac{1-b}{\bar{b}-\underline{b}} \log\left(\frac{1-b}{\bar{b}-\underline{b}}\right).$$

Also,

$$\begin{aligned}
\sum_{t=1}^{T_b} w_t^2 &\approx \frac{1}{T\left(\bar{b}-\underline{b}\right)^2} \int_{\underline{b}}^{b} \left[\log\left(\frac{1-a}{1-\underline{b}}\right)\right]^2 da \\
&= \frac{1}{T\left(\bar{b}-\underline{b}\right)^2} \int_{\underline{b}}^{b} \left[\log\left(1-a\right)\right]^2 da - \frac{2\log(1-\underline{b})}{T\left(\bar{b}-\underline{b}\right)^2} \int_{\underline{b}}^{b} \log\left(1-a\right) da \\
&\quad + \frac{\left[\log(1-\underline{b})\right]^2 (b-\underline{b})}{T\left(\bar{b}-\underline{b}\right)^2},
\end{aligned}$$

and

$$\int_{\underline{b}}^{b} \log\left(1-a\right) da = -(1-b)\log(1-b) + (1-\underline{b})\log(1-\underline{b}) + \underline{b} - b,$$

$$\begin{aligned}
\int_{\underline{b}}^{b} \left[\log\left(1-a\right)\right]^2 da &= -(1-b)\left[\log(1-b)\right]^2 + 2(1-b)\log(1-b) + 2b \\
&\quad + (1-\underline{b})\left[\log(1-\underline{b})\right]^2 - 2(1-\underline{b})\log(1-\underline{b}) - 2\underline{b}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\sum_{t=1}^{T} w_t^2 &\approx \frac{1}{T\left(\bar{b}-\underline{b}\right)^2} \int_{\underline{b}}^{\bar{b}} \left[\log\left(\frac{1-a}{1-\underline{b}}\right)\right]^2 da + \frac{(1-\bar{b})}{T\left(\bar{b}-\underline{b}\right)^2} \left[\log\left(\frac{1-\bar{b}}{1-\underline{b}}\right)\right]^2 \\
&= \frac{1}{T\left(\bar{b}-\underline{b}\right)^2} \int_{\underline{b}}^{\bar{b}} \left[\log\left(1-a\right)\right]^2 da - \frac{2\log(1-\underline{b})}{T\left(\bar{b}-\underline{b}\right)^2} \int_{\underline{b}}^{\bar{b}} \left[\log\left(1-a\right)\right] da \\
&\quad + \frac{\left[\log(1-\underline{b})\right]^2 (\bar{b}-\underline{b})}{T\left(\bar{b}-\underline{b}\right)^2} + \frac{(1-\bar{b})}{T\left(\bar{b}-\underline{b}\right)^2} \left[\log\left(\frac{1-\bar{b}}{1-\underline{b}}\right)\right]^2.
\end{aligned}$$

The above expressions simplify considerably if we set $\underline{b} = 0$. We have

$$\sum_{t=1}^{T_b} w_t \approx \frac{b}{\bar{b}} + \frac{(1-b)\log(1-b)}{\bar{b}},$$

$$\sum_{t=1}^{T_b} w_t^2 \approx \frac{-(1-b)\left[\log(1-b)\right]^2 + 2(1-b)\log(1-b) + 2b}{T\bar{b}^2},$$

and

$$\sum_{t=1}^{T} w_t^2 \approx \frac{2\bar{b} + 2(1-\bar{b})\log(1-\bar{b})}{T\bar{b}^2}.$$

Using these results in (30), we have

$$\begin{aligned}
&\frac{\omega_x^2}{x_{T+1}^2}\left[E\left(e_{T+1}^2/\sigma_{(2)}^2\right) - 1\right] \approx \phi^2\left(\sum_{t=1}^{T_b} w_t\right)^2 + \left(q^2-1\right)\sum_{t=1}^{T_b} w_t^2 + \sum_{t=1}^{T} w_t^2 \\
&= \phi^2\left[\frac{b}{\bar{b}} + \frac{(1-b)\log(1-b)}{\bar{b}}\right]^2 \\
&\quad + \left(q^2-1\right)\left[\frac{-(1-b)\left[\log(1-b)\right]^2 + 2(1-b)\log(1-b) + 2b}{T\left(\bar{b}\right)^2}\right] \\
&\quad + \frac{2\bar{b} + 2(1-\bar{b})\log(1-\bar{b})}{T\bar{b}^2}.
\end{aligned}$$

In practice if we choose $\bar{b}$ to be very close to unity but not unity then $(1-\bar{b})\log(1-\bar{b}) \approx 0$ and $(1-\bar{b})[\log(1-\bar{b})]^2 \approx 0$ and the result in (49) follows.

B.6 MSFE of weights in Figure 4

Figure 6 shows the MSFE corresponding to the weights plotted in Figure 3. The horizontal axis gives the break point and the vertical axis the corresponding MSFE. It can be seen from Figure 3 that the robust optimal weights discount past observations more quickly than the optimal weights and the fitted ExpS weights. As a result, the MSFE of the robust optimal weights is lower if the break is more recent and is higher if the break is earlier in the sample. This difference is substantial for $T = 50$ and $\phi^2 = 0.1$ for larger ϕ or T the MSFE is very similar for the three weights.

B.7 Robust optimal weights with higher order terms

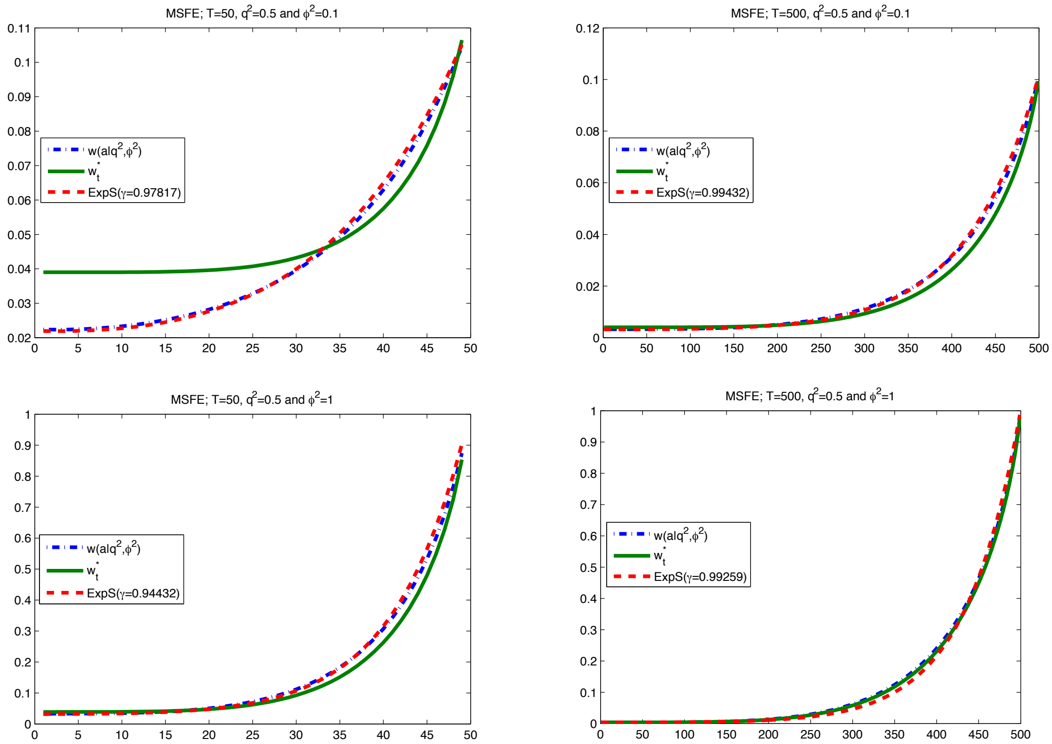
Consider now the second order term in (43) and let

$$T\phi^2 H(b,a) = -\frac{1}{(1-b)^2} + \frac{1}{b(1-b)^2} I(b-a),$$

and note that for $a < \underline{b}$

$$T\phi^2 \int_0^{\underline{b}} H(b,a) = 0 \text{ when } a < \underline{b},$$

Figure 6: MSFE of optimal weights, robust optimal weights, and fitted exponential smoothing weights in Figure 4



Note: The plots report the MSFE associated with the weights in Figure 4 for breaks at each point in the sample, $t = 1, 2, \ldots, T$. $T = 50$ in the plots in the left column, $T = 500$ in the plots in the right column, $\phi^2 = 0.1$ in the plots in the top row, $\phi^2 = 1$ in the plots in the bottom row, and $q^2 = 0.5$ throughout. The dash-dotted line represents the optimal weights $w(a|q^2, \phi^2)$ in (42), the solid line the robust optimal weights in (48), and the dashed line the ExpS weights in (10).

since by assumption the probability of drawing b less than $\underline{b}$ is zero. Consider now the value of the integral when $\underline{b} \leq a \leq \bar{b}$, and note that

$$\begin{aligned} T\phi^2 \int_{\underline{b}}^{\bar{b}} H(b,a) &= -\int_{\underline{b}}^{\bar{b}} \frac{1}{(1-b)^2} db + \int_{\underline{b}}^{\bar{b}} \frac{1}{b(1-b)^2} \mathrm{I}(b-a) db \\ &= -\int_{\underline{b}}^{a} \frac{1}{(1-b)^2} db - \int_{a}^{\bar{b}} \frac{1}{(1-b)^2} db + \int_{a}^{\bar{b}} \frac{1}{b(1-b)^2} db \\ &= -\int_{\underline{b}}^{a} \frac{1}{(1-b)^2} db + \int_{a}^{\bar{b}} \frac{1}{b(1-b)} db \\ &= -\frac{a-\underline{b}}{(1-a)(1-\underline{b})} + \log\left(\frac{\bar{b}}{\underline{b}}\right) + \log\left(\frac{1-\bar{b}}{1-a}\right). \end{aligned}$$

Finally, for $a > \bar{b}$ we have

$$T\phi^2 H(b,a) = -\int_{\underline{b}}^{\bar{b}} \frac{1}{(1-b)^2} db + \int_{\underline{b}}^{\bar{b}} \frac{1}{b(1-b)^2} \mathrm{I}(b-a) db = -\frac{\bar{b}-\underline{b}}{(1-\bar{b})(1-\underline{b})}.$$

Table 6: MSE of $\hat{\gamma}$ in Monte Carlo experiments with continuous breaks

$T\backslash\gamma$	0.8	0.9	0.95	0.98
50	0.013	0.012	0.009	0.004
100	0.005	0.003	0.005	0.002
200	0.003	0.002	0.002	0.001

Note: The table reports the MSE of the estimation of the MA parameter γ in an MA(1) model in the Monte Carlo experiments reported in Table 2.

Combining these results, we obtain

$$w(a) \approx \begin{cases} 0 & \text{for } a < \underline{b} \\ \frac{-1}{T(\bar{b}-\underline{b})} \log\left(\frac{1-a}{1-\underline{b}}\right) + \frac{1}{T^2\phi^2(\bar{b}-\underline{b})} \times & \\ \quad \times \left[\frac{-(a-\underline{b})}{(1-a)(1-\underline{b})} + \log\left(\frac{\bar{b}}{a}\right) + \log\left(\frac{1-\bar{b}}{1-a}\right)\right] & \text{for } \underline{b} \le a \le \bar{b} \\ \frac{-1}{T(\bar{b}-\underline{b})} \log\left(\frac{1-\bar{b}}{1-\underline{b}}\right) - \frac{1}{T^2\phi^2(\bar{b}-\underline{b})} \frac{\bar{b}-\underline{b}}{(1-\bar{b})(1-\underline{b})} & \text{for } a > \bar{b} \end{cases}$$

and the discrete time version is

$$w_t \approx \begin{cases} 0 & \text{for } t < T\underline{b} \\ \frac{-1}{T(\bar{b}-\underline{b})} \log\left(\frac{1-t/T}{1-\underline{b}}\right) + \frac{1}{T^2\phi^2(\bar{b}-\underline{b})} \times & \\ \quad \times \left[\frac{-[(t/T)-\underline{b}]}{(1-t/T)(1-\underline{b})} + \log\left(\frac{\bar{b}(1-\bar{b})}{(t/T)(1-t/T)}\right)\right] & \text{for } T\underline{b} \le t \le T\bar{b} \\ \frac{-1}{T(\bar{b}-\underline{b})} \log\left(\frac{1-\bar{b}}{1-\underline{b}}\right) - \frac{1}{T^2\phi^2} \frac{1}{(1-\bar{b})(1-\underline{b})} & \text{for } t > T\bar{b} \end{cases} \tag{63}$$

In the case where $\underline{b} = 0$, and $\bar{b}T = T - 1$, or $\bar{b} = 1 - 1/T$ we have for $1 \le t \le T - 1$

$$w_t^* = \frac{-1}{T-1} \log(1 - t/T) - \frac{1}{T(T-1)\phi^2} \left[\frac{t}{T-t} - \log\left(\frac{(T-1)}{t(T-t)}\right)\right], \tag{64}$$

and for the final date using the last part of (63) we obtain

$$w_T^* = \frac{\log(T)}{T-1} - \frac{1}{T\phi^2}. \tag{65}$$

The scaled version of these weights (that sum up to unity) are given by

$$w_t = \frac{w_t^*}{\sum_{s=1}^T w_s^*}, \text{ for } t = 1, 2, \ldots, T.$$

In practice, one could set $\phi^2 = 1/2$ or 1.

B.8 Additional results for the Monte Carlo experiments in the paper

The results in Table 7 show that the influence of a break in the error variance is of negligible importance for the forecasts, which confirms our theoretical results.

Table 7: Monte Carlo results for random walk model with a discrete break, $q = 0.5$

b	0.95			0.9		
λ	0.5	1	2	0.5	1	2
$T = 50$						
opt.weight(disc.break; b, λ)	0.927	0.656	0.284	0.915	0.637	0.277
estim.opt.weight(disc.break; $\hat{b}, \hat{\lambda}$)	1.042	0.853	0.411	1.048	0.835	0.333
rob.opt.weights($\underline{b} = 0.75, \bar{b} = 0.98$)	0.942	0.778	0.602	0.925	0.715	0.479
rob.opt.weights($\underline{b} = 0, \bar{b} = 1$)	0.955	0.856	0.750	0.940	0.810	0.662
post-break obs.($\hat{b}$)	1.065	0.864	0.410	1.073	0.850	0.334
opt.window($\hat{b}, \hat{\lambda}$)	1.002	0.824	0.426	1.006	0.803	0.335
AveW($w_{\min} = 0.05$)	0.964	0.887	0.804	0.947	0.835	0.709
estim.opt.weight(cont.break; $\hat{\delta}$)	0.997	0.939	0.648	0.999	0.895	0.424
ExpS($\hat{\gamma}$)	0.997	0.939	0.648	0.999	0.895	0.424
$T = 100$						
opt.weight(disc.break; b, λ)	0.896	0.605	0.257	0.878	0.593	0.257
estim.opt.weight(disc.break; $\hat{b}, \hat{\lambda}$)	1.031	0.807	0.311	1.021	0.731	0.262
rob.opt.weights($\underline{b} = 0.75, \bar{b} = 0.98$)	0.930	0.794	0.647	0.900	0.704	0.480
rob.opt.weights($\underline{b} = 0, \bar{b} = 1$)	0.952	0.866	0.774	0.931	0.805	0.662
post-break obs.($\hat{b}$)	1.054	0.821	0.311	1.042	0.742	0.261
opt.window($\hat{b}, \hat{\lambda}$)	0.994	0.780	0.316	0.990	0.714	0.265
AveW($w_{\min} = 0.05$)	0.964	0.899	0.830	0.939	0.830	0.706
estim.opt.weight(cont.break; $\hat{\delta}$)	0.990	0.910	0.469	0.978	0.769	0.289
ExpS($\hat{\gamma}$)	0.990	0.910	0.469	0.978	0.769	0.289
$T = 200$						
opt.weight(disc.break; b, λ)	0.871	0.572	0.238	0.863	0.578	0.248
estim.opt.weight(disc.break; $\hat{b}, \hat{\lambda}$)	1.024	0.704	0.243	0.992	0.613	0.249
rob.opt.weights($\underline{b} = 0.75, \bar{b} = 0.98$)	0.923	0.787	0.642	0.893	0.697	0.474
rob.opt.weights($\underline{b} = 0, \bar{b} = 1$)	0.949	0.863	0.771	0.928	0.802	0.658
post-break obs.($\hat{b}$)	1.047	0.714	0.243	1.010	0.616	0.248
opt.window($\hat{b}, \hat{\lambda}$)	0.991	0.686	0.245	0.971	0.607	0.249
AveW($w_{\min} = 0.05$)	0.962	0.898	0.831	0.937	0.828	0.704
estim.opt.weight(cont.break; $\hat{\delta}$)	0.979	0.800	0.288	0.952	0.648	0.263
ExpS($\hat{\gamma}$)	0.979	0.800	0.288	0.952	0.648	0.263

Note: Here, $q = \sigma_{(1)}/\sigma_{(2)} = 0.5$. Otherwise see footnote of Tables 2 and 3.

B.9 Monte Carlo results for an AR(1) DGP

Table 8 reports the results of a Monte Carlo experiment where we use the different forecasting methods outlined in the paper for an AR(1) model. While we cannot claim that the weights that we have derived are optimal in the context of this model, it is nonetheless interesting to assess their performance in a dynamic setting. Here, we use an AR(1)

$$y_t^{(r)} = \mu_t + \rho_t y_{t-1}^{(r)} + \varepsilon_t^{(r)}, \quad \varepsilon_t^{(r)} \sim \mathrm{N}(0, 1)$$

and set $\mu_t = 0$ and $\rho_t = 0.1$. If there is a break in μ_t we set $\mu_{(1)} = 0$ and $\mu_{(2)} = 1$ and if there is a break in ρ_t we wet $\rho_{(1)} = 0.1$ and $\rho_{(2)} = 0.6$. We generate data for $t = -99, \ldots, 0, 1, \ldots, T$ and discard the first 100 observations. The results are based on $r = 1, 2, \ldots, 10000$ repetitions.

Table 8 reports the MSFE relative to that of the equal weights forecast. We compare the optimal and robust optimal weights to those of the forecast based on the post break sample and the AveW sample. We omit the constant gain least square and the optimal window for computational simplicity and as these experiments only intend to give a first impression of the performance of the weights derived in this paper.

The results suggest that, if there is a break in ρ only and the unconditional mean remains zero, the "optimal" weights increase the MSFE above that of the equal weights forecast. The robust optimal weights, in contrast, lead to substantial improvements in the MSFE, also exceeding that of the AveW procedure.

If the intercept changes but the persistence of the process remains unchanged, the optimal weights generally perform better than the equal weights forecast. The robust optimal weights that use the last quarter of the observations perform best. The post break forecast and the AveW forecast improve over the full sample forecasts but not as much as the robust optimal weights.

Finally, when both the intercept and the AR(1) parameter change, the optimal weights lead to large and often the largest improvements in MSFE. Only in the smallest sample do the robust optimal weights still perform better. Overall, the results suggest that the optimal and the robust optimal weights do lead to large improvements in forecast performance also in dynamic models.

Table 8: Monte Carlo results for an AR(1) and a discrete break, $q = 1$

b	0.95			0.9		
λ_μ	0	1	1	0	1	1
λ_ρ	0.5	0	0.5	0.5	0	0.5
$T = 50$						
estim.opt.weight$(\hat{b}, \hat{\lambda}_\mu, \hat{\lambda}_\rho)$	1.043	1.044	0.885	1.094	0.979	0.675
rob.opt.weights$(\underline{b} = 0.75, \bar{b} = 0.98)$	0.928	0.878	0.540	0.937	0.842	0.501
rob.opt.weights$(\underline{b} = 0, \bar{b} = 1)$	0.895	0.873	0.599	0.898	0.855	0.547
post-break obs.$(\hat{b})$	1.097	1.089	0.918	1.173	1.029	0.744
AveW$(w_{\min} = 0.05)$	0.937	0.908	0.721	0.936	0.875	0.656
$T = 100$						
estim.opt.weight$(\hat{b}, \hat{\lambda}_\mu, \hat{\lambda}_\rho)$	1.070	0.946	0.528	1.104	0.866	0.445
rob.opt.weights$(\underline{b} = 0.75, \bar{b} = 0.98)$	0.893	0.797	0.448	0.887	0.755	0.445
rob.opt.weights$(\underline{b} = 0, \bar{b} = 1)$	0.887	0.836	0.522	0.882	0.809	0.504
post-break obs.$(\hat{b})$	1.140	0.994	0.553	1.190	0.902	0.494
AveW$(w_{\min} = 0.05)$	0.936	0.888	0.691	0.920	0.836	0.629
$T = 200$						
estim.opt.weight$(\hat{b}, \hat{\lambda}_\mu, \hat{\lambda}_\rho)$	1.055	0.817	0.333	1.025	0.734	0.398
rob.opt.weights$(\underline{b} = 0.75, \bar{b} = 0.98)$	0.870	0.774	0.397	0.851	0.742	0.443
rob.opt.weights$(\underline{b} = 0, \bar{b} = 1)$	0.881	0.827	0.473	0.867	0.804	0.503
post-break obs.$(\hat{b})$	1.127	0.854	0.362	1.078	0.751	0.421
AveW$(w_{\min} = 0.05)$	0.933	0.882	0.659	0.909	0.834	0.624

Note: The results are for an AR(1) model, $y_t^{(r)} = \mu_t + \rho_t y_{t-1}^{(r)} + \varepsilon_t^{(r)}$, $\varepsilon_t^{(r)} \sim \mathrm{N}(0, 1)$ with a single break in μ_t and/or in ρ_t at T_b. Furthermore, $\lambda_\mu = (\mu_{(1)} - \mu_{(0)})/\sigma$, $\lambda_\rho = (\rho_{(2)} - \rho_{(1)})/\sigma$, and here $\sigma = 1$. The results are based on 10,000 repetitions. For definitions and forecasting procedures see the footnotes of Tables 2 and 3.

B.10 Constant gain least squares

As stated in (51) and reproduced here for the convenience of the reader, in the linear regression model (1) the parameter is estimated using the recursion (Evans and Honkapohja, 2001)

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \alpha \mathbf{R}_{t-1}^{-1} \mathbf{x}_t (y_t - \mathbf{x}_t' \boldsymbol{\beta}_{t-1})$$

and

$$\mathbf{R}_t = \mathbf{R}_{t-1} + \alpha(\mathbf{x}_t \mathbf{x}_t' - \mathbf{R}_{t-1}).$$

We set $\boldsymbol{\beta}_0$ to the OLS estimate based on the first $2k$ observations and $\mathbf{R}_0 = \mathbf{I}_k$. In the empirical application we estimate α by minimizing the MSFE of pseudo-out-of-sample forecasts of the last 40 pre-sample observations, that is, for the period 1984Q1–1994Q4. The respective estimates of α are reported in Table 9.

Table 9: Estimates of the constant gain least squares updating parameter, $\hat{\alpha}$

Country	$h=1$	$h=2$	$h=3$	$h=4$
USA	0.03	0.21	0.26	0.20
Japan	0.09	0.20	0.22	0.34
Germany	0.04	0.17	0.19	0.29
UK	0.20	0.17	0.17	0.20
F	0.04	0.05	0.10	0.10
It.	0.07	0.10	0.11	0.32
Spain	0.22	0.24	0.27	0.27
Can.	0.11	0.26	0.23	0.23
Aus.	0.01	0.21	0.22	0.17

The table reports the estimated α for the CGLS estimation in the application. The estimates were obtained by minimizing the MSFE of pseudo-out-of-sample forecasts of the last 40 pre-sample observations.

C Country specific results from the application to the yield curve as predictor of real economic activity

Tables 11 reports the results for the first sub-sample of forecasts, 1994Q1–2000Q4.

The forecasts for the second sub-sample, 2001Q1–2006Q4 are in Table 12.

The results for the last sub-sample, 2007Q1–2009Q4, are in Table 13.

References

Chu, Ka Lok, Simo Puntanen, and George P. H. Styan (2009) 'Solution to problem 1/SP09: Inverse and determinant of a special symmetric matrix.' *Statistical Papers* 52, 258–260.

Evans, George W., and Seppo Honkapohja (2001) *Learning and Expectations in Macroeconomics* Princeton: Princeton University Press.

Neudecker, Heinz, Götz Trenkler, and Shuangzhe Liu (2009) 'Problem 1/SP09: Inverse and determinant of a special symmetric matrix.' *Statistical Papers* 50, 221.

Table 10: Predictive power of the yield curve: Relative forecast accuracy per country (all forecasts: 1994Q1–2009Q4)

	USA	Japan	Ger.	UK	F	It.	Spain	Can.	Aus.
$h = 1$									
prop. breaks	0.063	1.000	0.250	1.000	0.266	1.000	0.094	0.266	1.000
equal weight(MSFE)	0.463	1.097	0.682	0.469	0.293	0.542	0.343	0.444	0.350
est.asy.opt.weight	0.918	0.934	0.969	0.836	0.966	1.249	0.928	0.977	0.986
rob.opt.weight(1 break)	0.877	0.874	0.941	0.853	0.957	0.914	0.915	0.905	0.999
rob.opt.weight(2 breaks)	0.934	0.922	0.964	0.930	0.979	0.952	0.954	0.938	1.002
post-break	1.046	0.888	1.289	0.962	1.484	1.014	1.794	1.278	0.969
AveW	1.001	0.982	0.980	0.985	0.996	0.986	0.998	1.003	0.985
CGLS($\hat{\alpha}$)	0.904	0.764	0.952	16.502	0.967	0.820	11.488	0.837	6.269
CGLS($\bar{\hat{\alpha}}$)	0.854	0.765	0.919	1.075	0.908	0.800	1.003	0.858	1.137
$h = 2$									
prop. breaks	0.095	0.984	0.524	1.000	0.286	1.000	0.000	0.175	0.000
equal weight(MSFE)	1.533	3.139	1.871	1.635	0.874	1.627	1.248	1.615	0.724
est.asy.opt.weight	0.947	1.014	1.116	0.973	1.131	1.360	1.000	1.195	1.000
rob.opt.weight(1 break)	0.852	0.849	0.997	0.947	1.000	0.913	0.988	0.906	1.058
rob.opt.weight(2 breaks)	0.923	0.904	0.977	0.977	0.995	0.947	0.988	0.932	1.025
post-break	1.137	0.965	1.247	1.012	1.238	1.023	1.000	1.767	1.000
AveW	1.007	0.981	0.995	1.003	0.999	0.983	0.999	1.003	0.993
CGLS($\hat{\alpha}$)	1.005	0.714	1.182	2.614	1.004	0.809	2.494	1.463	1.822
CGLS($\bar{\hat{\alpha}}$)	0.927	0.710	1.189	3.184	1.201	0.905	1.466	1.158	1.790
$h = 3$									
prop. breaks	0.177	0.710	0.516	0.758	0.113	1.000	0.000	0.258	0.000
equal weight(MSFE)	3.059	5.751	3.473	3.240	1.722	3.146	2.598	3.284	1.144
est.asy.opt.weight	0.975	1.026	1.415	1.023	1.081	1.132	1.000	1.126	1.000
rob.opt.weight(1 break)	0.896	0.844	1.041	1.003	1.035	0.938	1.009	0.921	1.066
rob.opt.weight(2 breaks)	0.942	0.894	0.987	1.002	1.009	0.958	1.007	0.933	1.030
post-break	1.094	0.981	1.250	1.018	1.647	1.035	1.000	1.171	1.000
AveW	1.001	0.979	1.000	1.004	1.001	0.993	1.000	1.003	0.996
CGLS($\hat{\alpha}$)	1.443	0.860	1.232	1.276	1.091	0.742	2.379	1.107	1.891
CGLS($\bar{\hat{\alpha}}$)	1.031	0.850	1.237	1.376	1.300	0.818	2.958	1.026	1.684
$h = 4$									
prop. breaks	0.049	0.361	0.508	0.754	0.311	0.016	0.000	0.033	0.754
equal weight(MSFE)	4.780	9.121	5.232	4.913	2.738	5.029	4.200	5.095	1.513
est.asy.opt.weight	1.012	1.033	1.180	1.054	1.215	0.948	1.000	1.062	1.115
rob.opt.weight(1 break)	0.931	0.834	1.074	1.029	1.063	0.980	1.019	0.951	1.058
rob.opt.weight(2 breaks)	0.953	0.883	0.990	1.011	1.021	0.977	1.019	0.943	1.025
post-break	1.027	1.049	1.262	1.039	1.281	0.953	1.000	1.065	0.989
AveW	0.996	0.978	1.007	1.010	1.004	1.010	1.001	1.002	0.990
CGLS($\hat{\alpha}$)	1.025	1.647	1.324	1.377	1.138	0.817	3.587	1.215	1.429
CGLS($\bar{\hat{\alpha}}$)	1.022	0.994	1.280	1.560	1.370	0.759	8.399	1.207	1.627

Note: The line denoted "prop. break" reports the proportion of forecasts where a break was detected by the Bai and Perron (1997,2003) test. The countries are: USA, Japan, Germany, UK, France, Italy, Spain, Canada, and Australia. For the forecasting methods see footnote of Table 5.

Table 11: Predictive power of the yield curve: Relative forecast accuracy per country (subsample 1: 1994Q1–2000Q4)

	USA	Japan	Ger.	UK	F	It.	Spain	Can.	Aus.
$h = 1$									
prop. breaks	0.000	1.000	0.143	1.000	0.393	1.000	0.107	0.500	1.000
equal weight(MSFE)	0.317	0.834	0.404	0.156	0.187	0.300	0.199	0.292	0.482
est.asy.opt.weight	1.000	1.009	1.028	0.934	1.030	1.036	0.985	1.162	0.983
rob.opt.weight(1 break)	0.794	0.800	0.955	0.773	1.020	1.116	0.787	0.978	0.979
rob.opt.weight(2 breaks)	0.904	0.861	0.985	0.935	1.011	1.024	0.915	0.974	0.996
post-break	1.000	0.938	0.997	0.729	1.444	1.216	1.021	1.246	0.949
AveW	1.003	0.970	0.950	1.003	1.000	1.041	0.977	1.016	0.978
CGLS($\hat{\alpha}$)	0.874	0.738	0.990	0.711	1.057	1.098	1.307	0.810	3.806
CGLS($\bar{\hat{\alpha}}$)	0.768	0.738	0.984	0.738	0.994	1.108	0.830	0.871	1.009
$h = 2$									
prop. breaks	0.111	0.963	0.333	1.000	0.333	1.000	0.000	0.185	0.000
equal weight(MSFE)	0.962	2.280	0.813	0.561	0.534	0.504	0.635	1.088	0.792
est.asy.opt.weight	0.789	0.972	1.461	0.719	1.280	1.470	1.000	1.031	1.000
rob.opt.weight(1 break)	0.611	0.721	1.022	0.777	1.121	1.292	0.788	0.981	1.010
rob.opt.weight(2 breaks)	0.827	0.806	0.993	0.976	1.057	1.046	0.940	0.951	1.006
post-break	0.772	1.009	2.138	0.734	1.353	1.512	1.000	1.151	1.000
AveW	1.028	0.968	0.983	0.980	1.002	1.136	0.978	1.020	0.987
CGLS($\hat{\alpha}$)	0.499	0.686	1.153	0.731	1.133	1.259	4.542	0.753	1.230
CGLS($\bar{\hat{\alpha}}$)	0.469	0.633	1.156	0.731	1.218	1.541	1.655	0.739	1.208
$h = 3$									
prop. breaks	0.385	0.423	0.385	0.500	0.115	1.000	0.000	0.538	0.000
equal weight(MSFE)	2.090	4.085	1.566	1.133	1.150	0.779	1.239	2.546	1.193
est.asy.opt.weight	0.701	0.976	3.300	0.846	1.042	2.064	1.000	1.170	1.000
rob.opt.weight(1 break)	0.552	0.690	1.084	0.790	1.158	1.444	0.792	0.990	0.982
rob.opt.weight(2 breaks)	0.819	0.768	1.010	1.001	1.070	1.070	0.967	0.937	0.993
post-break	0.692	1.014	2.362	0.937	1.101	1.713	1.000	1.236	1.000
AveW	1.003	0.963	0.997	0.972	1.004	1.283	0.980	1.019	0.990
CGLS($\hat{\alpha}$)	0.436	0.758	1.490	0.701	1.106	1.407	5.065	0.829	1.224
CGLS($\bar{\hat{\alpha}}$)	0.345	0.725	1.505	0.687	1.262	2.121	8.930	0.825	1.180
$h = 4$									
prop. breaks	0.120	0.560	0.320	0.440	0.560	0.000	0.000	0.040	0.400
equal weight(MSFE)	3.744	7.274	2.453	1.839	1.976	1.284	2.023	4.490	1.650
est.asy.opt.weight	1.036	1.062	1.369	0.926	1.025	1.000	1.000	1.008	1.204
rob.opt.weight(1 break)	0.570	0.697	1.147	0.820	1.184	1.506	0.819	1.019	0.922
rob.opt.weight(2 breaks)	0.836	0.767	1.021	1.025	1.085	1.087	0.998	0.943	0.965
post-break	1.085	1.155	2.360	0.995	1.078	1.000	1.000	1.000	0.971
AveW	0.974	0.969	1.018	0.978	1.006	1.331	0.981	1.017	0.979
CGLS($\hat{\alpha}$)	0.285	1.184	1.740	0.668	1.132	3.435	13.700	0.886	1.012
CGLS($\bar{\hat{\alpha}}$)	0.281	0.991	1.507	0.774	1.334	2.376	38.320	0.883	1.073

Note: See footnote of Table 5. The dates given above denote the periods for which one-period ahead forecasts are made. The $h = 2$ forecast makes the first forecast for the observation one quarter later, the $h = 3$ forecast for that two periods later, and the $h = 4$ forecast for that three quarters later.

Table 12: Predictive power of the yield curve: Relative forecast accuracy per country (subsample 2: 2001Q1–2006Q4)

	USA	Japan	Ger.	UK	F	It.	Spain	Can.	Aus.
$h = 1$									
prop. breaks	0.000	1.000	0.333	1.000	0.083	1.000	0.000	0.000	1.000
equal weight(MSFE)	0.227	0.430	0.265	0.072	0.139	0.181	0.070	0.197	0.199
est.asy.opt.weight	1.000	1.075	1.010	1.003	1.119	0.851	1.000	1.000	0.946
rob.opt.weight(1 break)	1.041	0.955	0.823	1.053	1.040	0.886	0.913	0.938	1.005
rob.opt.weight(2 breaks)	1.010	0.950	0.903	1.014	1.020	0.978	0.925	0.961	1.002
post-break	1.000	1.121	1.029	1.009	1.276	0.811	1.000	1.000	0.941
AveW	1.001	0.984	0.959	1.002	0.993	0.921	0.986	0.992	0.981
CGLS($\hat{\alpha}$)	1.019	0.982	0.852	1.314	1.054	0.950	0.955	1.138	10.894
CGLS($\bar{\hat{\alpha}}$)	1.133	0.982	0.754	1.191	1.113	0.941	1.072	1.102	1.168
$h = 2$									
prop. breaks	0.000	1.000	0.542	1.000	0.250	1.000	0.000	0.167	0.000
equal weight(MSFE)	0.614	1.181	0.849	0.145	0.281	0.497	0.185	0.681	0.495
est.asy.opt.weight	1.000	1.015	0.978	1.049	1.724	0.804	1.000	1.735	1.000
rob.opt.weight(1 break)	1.184	0.980	0.885	1.087	1.070	0.831	0.942	0.892	1.090
rob.opt.weight(2 breaks)	1.039	0.950	0.915	0.989	1.025	0.970	0.922	0.924	1.036
post-break	1.000	0.957	0.985	1.072	2.159	0.724	1.000	2.274	1.000
AveW	0.999	0.982	0.987	0.980	0.996	0.847	0.979	0.993	0.996
CGLS($\hat{\alpha}$)	1.952	1.289	0.732	1.717	1.106	0.911	1.279	2.495	3.139
CGLS($\bar{\hat{\alpha}}$)	1.516	1.291	0.727	1.723	1.594	0.832	1.397	1.574	2.400
$h = 3$									
prop. breaks	0.000	0.875	0.625	1.000	0.125	1.000	0.000	0.083	0.000
equal weight(MSFE)	1.096	1.905	1.711	0.253	0.460	0.831	0.355	1.290	0.843
est.asy.opt.weight	1.000	1.150	0.865	0.907	1.544	0.787	1.000	1.466	1.000
rob.opt.weight(1 break)	1.523	0.977	0.922	1.035	1.129	0.871	1.077	0.854	1.134
rob.opt.weight(2 breaks)	1.140	0.930	0.919	0.947	1.046	1.008	0.974	0.893	1.052
post-break	1.000	1.175	0.845	0.892	1.913	0.718	1.000	1.617	1.000
AveW	1.002	0.975	0.992	0.960	1.001	0.815	0.978	0.995	1.000
CGLS($\hat{\alpha}$)	5.363	2.096	1.014	1.963	1.707	1.010	3.825	2.403	2.244
CGLS($\bar{\hat{\alpha}}$)	2.713	1.961	1.013	2.054	2.937	0.922	2.272	2.185	1.958
$h = 4$									
prop. breaks	0.000	0.333	0.875	1.000	0.208	0.000	0.000	0.042	1.000
equal weight(MSFE)	1.569	2.457	2.645	0.373	0.660	1.084	0.568	1.989	1.020
est.asy.opt.weight	1.000	1.118	1.596	0.856	3.184	1.000	1.000	1.384	1.064
rob.opt.weight(1 break)	1.703	0.963	0.939	0.922	1.200	1.023	1.141	0.848	1.172
rob.opt.weight(2 breaks)	1.176	0.889	0.906	0.871	1.071	1.084	1.004	0.885	1.057
post-break	1.000	0.985	0.998	0.824	3.723	1.000	1.000	1.422	1.011
AveW	1.010	0.964	1.001	0.933	1.006	0.853	0.983	0.998	1.002
CGLS($\hat{\alpha}$)	3.704	5.951	1.459	2.371	1.990	1.313	4.366	2.908	1.804
CGLS($\bar{\hat{\alpha}}$)	3.749	2.405	1.447	2.530	4.473	1.553	3.289	2.868	2.437

Note: See footnote of Table 5. The dates given above denote the periods for which forecasts are made at all horizons.

Table 13: Predictive power of the yield curve: Relative forecast accuracy per country (subsample 3: 2007Q1-2009Q4)

	USA	Japan	Ger.	UK	F	It.	Spain	Can.	Aus.
$h = 1$									
prop. breaks	0.333	1.000	0.333	1.000	0.333	1.000	0.250	0.250	1.000
equal weight(MSFE)	1.277	3.046	2.168	1.997	0.850	1.831	1.227	1.294	0.346
est.asy.opt.weight	0.842	0.847	0.934	0.806	0.882	1.409	0.898	0.873	1.046
rob.opt.weight(1 break)	0.866	0.899	0.964	0.854	0.898	0.843	0.964	0.856	1.059
rob.opt.weight(2 breaks)	0.924	0.953	0.970	0.924	0.948	0.920	0.971	0.912	1.022
post-break	1.089	0.790	1.480	1.001	1.573	0.977	2.176	1.381	1.066
AveW	1.000	0.990	0.998	0.980	0.995	0.977	1.007	1.000	1.009
CGLS($\hat{\alpha}$)	0.881	0.719	0.960	20.462	0.893	0.689	16.537	0.760	8.942
CGLS($\bar{\hat{\alpha}}$)	0.804	0.721	0.932	1.128	0.796	0.655	1.061	0.777	1.517
$h = 2$									
prop. breaks	0.250	1.000	0.917	1.000	0.250	1.000	0.000	0.167	0.000
equal weight(MSFE)	4.656	8.984	6.297	7.033	2.825	6.414	4.752	4.669	1.031
est.asy.opt.weight	1.006	1.038	1.054	1.015	0.950	1.427	1.000	1.123	1.000
rob.opt.weight(1 break)	0.877	0.888	1.020	0.972	0.934	0.858	1.052	0.872	1.110
rob.opt.weight(2 breaks)	0.936	0.948	0.989	0.976	0.963	0.926	1.008	0.925	1.049
post-break	1.343	0.942	1.059	1.060	1.006	0.983	1.000	1.942	1.000
AveW	0.999	0.988	1.001	1.008	0.998	0.977	1.007	0.998	1.003
CGLS($\hat{\alpha}$)	0.991	0.579	1.312	2.990	0.929	0.714	1.973	1.534	1.580
CGLS($\bar{\hat{\alpha}}$)	0.984	0.601	1.322	3.685	1.116	0.804	1.415	1.256	2.210
$h = 3$									
prop. breaks	0.083	1.000	0.583	0.833	0.083	1.000	0.000	0.000	0.000
equal weight(MSFE)	9.086	17.054	11.129	13.781	5.487	12.903	10.032	8.870	1.640
est.asy.opt.weight	1.106	1.024	1.009	1.059	1.020	1.055	1.000	1.000	1.000
rob.opt.weight(1 break)	0.916	0.895	1.064	1.039	0.963	0.881	1.062	0.898	1.129
rob.opt.weight(2 breaks)	0.955	0.951	1.001	1.004	0.974	0.937	1.020	0.942	1.066
post-break	1.317	0.921	1.035	1.037	1.850	0.987	1.000	1.000	1.000
AveW	1.000	0.987	1.004	1.011	1.001	0.978	1.007	0.995	1.000
CGLS($\hat{\alpha}$)	0.999	0.636	1.221	1.354	0.982	0.620	1.559	0.903	2.580
CGLS($\bar{\hat{\alpha}}$)	0.968	0.667	1.225	1.474	1.042	0.634	1.409	0.815	2.197
$h = 4$									
prop. breaks	0.000	0.000	0.167	0.917	0.000	0.083	0.000	0.000	1.000
equal weight(MSFE)	13.360	26.297	16.192	20.399	8.480	20.720	15.997	12.569	2.214
est.asy.opt.weight	1.000	1.000	0.984	1.085	1.000	0.936	1.000	1.000	1.023
rob.opt.weight(1 break)	0.961	0.889	1.095	1.072	0.983	0.908	1.063	0.932	1.163
rob.opt.weight(2 breaks)	0.969	0.948	1.008	1.013	0.982	0.951	1.025	0.960	1.087
post-break	1.000	1.000	1.002	1.056	1.000	0.942	1.000	1.000	0.995
AveW	1.005	0.987	1.005	1.019	1.002	0.984	1.007	0.992	0.998
CGLS($\hat{\alpha}$)	0.828	1.109	1.149	1.473	1.009	0.428	0.867	0.925	1.731
CGLS($\bar{\hat{\alpha}}$)	0.815	0.733	1.155	1.672	0.904	0.467	0.879	0.922	1.741

Note: See footnote of Table 5. The dates given above denote the periods for which forecasts are made at all horizons.