# Appendix A: Asymptotic Properties of the GMM Estimators and Diagnostic Tests

## 7.1 Proof of Theorem 1

The additional regularity conditions required for consistency are stated in Hansen's Theorem 2.1. It follows from assumption 3 that $a_T^* \xrightarrow{a.s.} a_0^* = D_0' W_0^*$. Define $h_0^*(b^*) = a_0^*[E(R^e) - D_0 b^*]$. Given that $D_0$ has full column rank and $W_0^*$ is positive definite, the function $h_0^*(b^*)$ has a unique zero, $b_z^* = \left(D_0' W_0^* D_0\right)^{-1} D_0' W_0^* E(R^e)$. Since the model is true $E(R^e) = D_0 b_0^*$. Substituting this into the expression for $b_z^*$ we have $b_z^* = b_0^*$. From Hansen's (1982) Theorem 2.1, $\hat{b}^* \xrightarrow{a.s.} b_0^*$.

Similarly, it follows from assumption 3 that $a_T^\diamond \xrightarrow{a.s.} a_0^\diamond$ with

$$a_0^\diamond = \begin{pmatrix} d_0' W_0^\diamond & 0 \\ 0 & I_k \end{pmatrix}.$$

Define

$$h_0^\diamond(b^\diamond, \mu) = a_0^\diamond \begin{pmatrix} E(R^e) - d_0 b^\diamond \\ E(f_t) - \mu \end{pmatrix}.$$

Given that $d_0$ has full column rank and $W_0^\diamond$ is positive definite, the function $h_0^\diamond(b^\diamond, \mu)$ has a unique zero, $b_z^\diamond = \left(d_0' W_0^\diamond d_0\right)^{-1} d_0' W_0^\diamond E(R^e)$, $\mu_z = E(f_t) = \mu_0$. Since the model is true $E(R^e) = d_0 b_0^\diamond$. Substituting this into the expression for $b_z^\diamond$ we have $b_z^\diamond = b_0^\diamond$. From Hansen's (1982) Theorem 2.1, $\hat{b}^\diamond \xrightarrow{a.s.} b_0^\diamond$.

The matrices $\delta_0^*$ and $\delta_0^\diamond$ have full column rank due to the properties of $D_0$ and $d_0$. It follows from Hansen's Theorem 3.1 that $\hat{b}^*$ and $\hat{\theta}^\diamond$ have the asymptotic distributions stated in the theorem.

The model-predicted expected returns are $D_T \hat{b}^*$ and $d_T \hat{b}^\diamond$, respectively, for the two normalizations. Given that results above these both converge almost surely to $E(R_t^e)$ and, therefore, we get the result that $R_*^2 \xrightarrow{a.s.} 1$ and $R_\diamond^2 \xrightarrow{a.s.} 1$.

From the results above it follows that $\delta_T^* \xrightarrow{a.s.} \delta_0^*$, $\delta_T^\diamond \xrightarrow{a.s.} \delta_0^\diamond$, $S_T^* \xrightarrow{a.s.} S_0^*$ and $S_T^\diamond \xrightarrow{a.s.} S_0^\diamond$. Also $A_T^* \xrightarrow{a.s.} A_0^* = I_n - \delta_0^*(a_0^* \delta_0^*)^{-1} a_0^*$ and $A_T^\diamond \xrightarrow{a.s.} A_0^\diamond = I_{n+k} - \delta_0^\diamond(a_0^\diamond \delta_0^\diamond)^{-1} a_0^\diamond$. The results concerning the asymptotic distributions of $J^*$ and $J^\diamond$ follow from Hansen's Lemma 4.1.■

## 7.2 Proof of Theorem 2

Since $D_0$ has rank less than $k$, the function $h_0^*(b^*) = a_0^*[E(R^e) - D_0 b^*]$, defined in the proof of Theorem 1, does not have a unique zero. Instead any $b^*$ such that $\left(D_0' W_0^* D_0\right) b^* = D_0' W_0^* D_0 b_0^*$

is a zero of $h_0^*(b^*)$. This means that $b_0^* + x$ is a zero for any $x$ in the nullspace of $D_0' W_0^* D_0$–denoted $\mathcal{N}(D_0' W_0^* D_0)$—which is a non-empty set when $\text{rank}(D_0) < k$. So $b^*$ is asymptotically unidentified.

As in the proof of Theorem 1, the last $k$ rows of the function $h_0^\diamond(b^\diamond, \mu)$ have a unique zero, $\mu_z = E(f_t) = \mu_0$. However, because $d_0$ has rank less than $k$, the first $n$ rows of the function $h_0^\diamond(b^\diamond, \mu)$, which are $d_0' W_0^\diamond [E(R^e) - d_0 b^\diamond]$, do not have a unique zero. Instead any $b^\diamond$ such that $(d_0' W_0^\diamond d_0)\, b^\diamond = d_0' W_0^\diamond d_0 b_0^\diamond$ is a zero. This means that $b_0^\diamond + x$ is a zero for any $x \in \mathcal{N}(d_0' W_0^\diamond d_0)$, which is a non-empty set when $\text{rank}(d_0) < k$. So $b^\diamond$ is asymptotically unidentified.

The predicted expected returns from the $a$-normalization are $D_T \hat{b}^*$. Although $\hat{b}^*$ is not uniquely identified asymptotically, it lies almost surely in the set $B_0^* = \{b | b - b_0^* = x, x \in \mathcal{N}(D_0' W_0^* D_0)\}$. Since $W_0^*$ is positive definite, any $x \in \mathcal{N}(D_0' W_0^* D_0)$ is in $\mathcal{N}(D_0)$. Therefore $D_T \hat{b}^* \stackrel{a.s.}{\to} D_0 b_0^* = E(R_t^e)$. Therefore $R_*^2 \stackrel{a.s.}{\to} 1$. A similar result holds for $R_\diamond^2$.■

## 7.3 Proof of Theorem 3

As in the proof to Theorem 1, $a_T^* \stackrel{a.s.}{\to} a_0^* = D_0' W_0^*$. Because $D_0$ has full column rank and $W_0^*$ is positive definite, the function $h_0^*(b^*)$ has a unique zero, $b_s^* = (D_0' W_0^* D_0)^{-1} D_0' W_0^* E(R^e)$. From Hansen's (1982) Theorem 2.1, $\hat{b}^* \stackrel{a.s.}{\to} b_s^*$. Of course, since the model is false, $b_s^*$ does not have an interpretation as a "true" parameter value.

To get the expression for $b_s^*$ in the statement of the theorem proceed as follows. Let $x$ be the unique element of $\mathcal{N}(d_0)$ whose elements sum to 1 (all other elements of $\mathcal{N}(d_0)$ are proportional to $x$ because $d_0$ has rank $k-1$). Let $X = (\ X_1 \quad x\ )$ where $X_1$ is a $k \times (k-1)$ matrix whose columns span the rowspace of $d_0$, denoted $\mathcal{R}(d_0) = \mathcal{N}(d_0)^\perp$. The columns of $X$ span $R^k$, by construction. Define $\tilde{b}_s^* = X^{-1} b_s^*$ and let $\tilde{b}_{s1}^*$ denote the first $k-1$ elements of $\tilde{b}_s^*$ and $\tilde{b}_{sk}^*$ denote the $k$th element of $\tilde{b}_s^*$. It follows that

$$\begin{aligned} E(R^e) - D_0 b_s^* &= E(R^e) - D_0 X X^{-1} b_s^* \\ &= E(R^e) - D_0 X \tilde{b}_s^* \\ &= E(R^e) - D_0 X_1 \tilde{b}_{s1}^* - D_0 x \tilde{b}_{sk}^* \\ &= E(R^e) - D_0 X_1 \tilde{b}_{s1}^* - [d_0 + E(R^e) E(f')] x \tilde{b}_{sk}^*. \end{aligned}$$

Since $x \in \mathcal{N}(d_0)$, $d_0 x = 0$, so we can write

$$E(R^e) - D_0 b_s^* = E(R^e) \left[1 - E(f') x \tilde{b}_{sk}^*\right] - D_0 X_1 \tilde{b}_{s1}^*.$$

This means we can set $E(R^e) - D_0 b_s^* = 0$ by choosing $\tilde{b}_{s1}^* = 0$ and $\tilde{b}_{sk}^* = 1/[E(f)'x]$. Since $b_s^* = X\tilde{b}_s^*$ it follows that $b_s^* = x/[x'E(f_t)]$. By assumption $x'E(f_t)$ cannot be zero, otherwise $\text{rank}[E(R_t^e f_t')] < k$ and we also know that at least one element of $x$ is non-zero, so this means at least one element of $b_s^*$ is non-zero. Since $E(R^e) = D_0 b_s^*$ we also have $R_*^2 \overset{a.s.}{\to} 1$.

As in the proof of Theorem 1, the last $k$ rows of the function $h_0^\diamond(b^\diamond, \mu)$ have a unique zero, $\mu_z = E(f_t) = \mu_0$. However, because $d_0$ has rank less than $k$, the first $n$ rows of the function $h_0^\diamond(b^\diamond, \mu)$, which are $d_0' W_0^\diamond [E(R^e) - d_0 b^\diamond]$, do not have a unique zero. Instead any $b^\diamond$ such that $(d_0' W_0^\diamond d_0)\, b^\diamond = d_0' W_0^\diamond E(R^e)$ is a zero. Let $b_z^\diamond$ be a zero. This means that $b_z^\diamond + x$ is a zero for any $x$ in the nullspace of $d_0' W_0^\diamond d_0$, which is a non-empty set because $\text{rank}(d_0) < k$. So $b^\diamond$ is asymptotically unidentified. Although there are arbitrarily many solutions to $d_0' W_0^\diamond [E(R^e) - d_0 b^\diamond] = 0$, in general, there is no solution to $E(R^e) - d_0 b^\diamond = 0$.■

## 7.4 Proof of Theorem 4

The matrix $\delta_0^* = D_0$ has full column rank. It follows from Hansen's Theorem 3.1 that $\hat{b}^*$ has the asymptotic distribution stated in the theorem.

From the results above it follows that $\delta_T^* \overset{a.s.}{\to} \delta_0^*$ and $A_T^* \overset{a.s.}{\to} A_0^* = I_n - \delta_0^*(a_0^*\delta_0^*)^{-1}a_0^*$, however, the matrix $S_T^*$ will not generally be a consistent estimator for $S_s^*$ because it imposes the restriction that $E(u_t^* u_{t-j}^{*\prime}) = 0$ for $j \neq 0$. This restriction only holds when the model is true. Instead $S_T^* \overset{a.s.}{\to} V_s^* = E[u_t^*(b_s^*) u_t^*(b_s^*)']$.

This means that $\hat{V}_g^* \overset{a.s.}{\to} V_g^* = A_0^* V_s^* A_0^{*\prime}$. We also know $\sqrt{T} g_T^*(\hat{b}^*) \overset{d}{\to} N(0, V_{g0}^*)$ where $V_{g0}^* = A_0^* S_s^* A_0^{*\prime}$. Diagonalize $V_g^*$ as $V_g^* = P_g \Lambda_g P_g'$ where the columns of $P_g$ are the orthonormal eigenvectors of $V_g^*$ and $\Lambda_g$ is a diagonal matrix with the eigenvalues of $V_g^*$ on the diagonal. Diagonalize $V_{g0}^*$ as $V_{g0}^* = P_0 \Lambda_0 P_0'$. Let $\tilde{\Lambda}_g = \Lambda_g^+$ and $\tilde{\Lambda}_0 = \Lambda_0^+$. These are diagonal matrices with zeros where $\Lambda_g$ and $\Lambda_0$ have zeros, and whose non-zero elements are the inverses of the non-zero elements of $\Lambda_g$ and $\Lambda_0$.

From these results it follows that $J^* \overset{d}{\to} Z'\Omega Z$, with $Z = \sqrt{T}\tilde{\Lambda}_0^{1/2} P_0' g_T^*(\hat{b}^*)$ and

$$\Omega = \Lambda_0^{1/2} P_0' P_g \tilde{\Lambda}_g P_g' P_0 \Lambda_0^{1/2}.$$

The vector $Z$ converges in distribution to a vector of independent normal random variables, the first $n-k$ of which have unit variance and the last $k$ of which have zero variance. The matrix $\Omega$ can be diagonalized as $\Omega = P_\Omega \Lambda_\Omega P_\Omega'$. When $V_s^* = S_s^*$ the first $n-k$ eigenvalues on the diagonal of $\Lambda_\Omega$ are ones while the rest are zeros. In this case $J \overset{d}{\to} \chi_{n-k}^2$. In general, however, $V_s^* \neq S_s^*$ and these eigenvalues will not be 1, so that $J^* \overset{d}{\to} \sum_{i=1}^{n-k} \lambda_{\Omega i} z_i^2$ where $\lambda_{\Omega 1}$,

$\lambda_{\Omega 2}$, ..., $\lambda_{\Omega n-k}$ are the nonzero eigenvalues of $\Omega$ and $z_1$, $z_2$, ..., $z_{n-k}$ are mutually independent standard normal random variables. Given the form of $\Omega$, $\prod_{i=1}^{n-k}\lambda_{\Omega i} = \prod_{i=1}^{n-k}\lambda_{0i}/\lambda_{gi}$, however, in general, $\lambda_{\Omega i} \neq \lambda_{0i}/\lambda_{gi}$.■

## 7.5 Proof of Theorem 5

Let $R_t^e = \mu_R + u_t$ with $E(u_t u_t') = \Sigma_R$, and $f_t = \mu + \epsilon_t$ with $E(\epsilon_t^2) = \sigma_f^2$. The asymptotic distribution of $\hat{b}^\diamond$ depends on the asymptotic distribution of $d_T = \frac{1}{T}\sum_{t=1}^T R_t^e\left(f_t - \bar{f}\right)$. Scaling $d_T$ by a factor of $T^{\frac{1}{2}}$ we have

$$T^{\frac{1}{2}} d_T = T^{-\frac{1}{2}}\sum_{t=1}^T u_t\epsilon_t - \sum_{t=1}^T \epsilon_t T^{-\frac{1}{2}}\sum_{t=1}^T u_t. \tag{37}$$

The first expression on the right hand side of (37) converges in distribution to $X \sim N(0, \sigma_f^2\Sigma_u)$. The second expression converges in probability to 0. So $T^{\frac{1}{2}} d_T \xrightarrow{d} X$. Also, $\bar{R}^e \xrightarrow{p} \mu_R$.

At the first stage of GMM the weighting matrix is $W_T^\diamond = I_n$ so we have $T^{-\frac{1}{2}}\hat{b}^\diamond = T^{\frac{1}{2}} d_T'\bar{R}^e/(T^{\frac{1}{2}} d_T' T^{\frac{1}{2}} d_T)$. It follows that $T^{-\frac{1}{2}}\hat{b}^\diamond \xrightarrow{d} Z = (X'\mu_R)/(X'X)$. The $t$-statistic for $\hat{b}^\diamond$ is $t = \hat{b}^\diamond/\sqrt{V_b^\diamond}$ where $V_b^\diamond$ is the first element on the diagonal of

$$V_\theta^\diamond = (a_T^\diamond\delta_T^\diamond)^{-1} a_T^\diamond S_T^\diamond a_T^{\diamond\prime}(\delta_T^{\diamond\prime} a_T^\diamond)^{-1}/T, \tag{38}$$

where $a_T^\diamond$ and $\delta_T^\diamond$ are defined in section 3.2 and $S_T^\diamond$ is a conventional estimate of the long-run covariance of the GMM errors in the first stage, which are

$$\begin{aligned} \hat{u}_{1t} &= R_t^e[1 - (f_t - \bar{f})\hat{b}^\diamond] \\ \hat{u}_{2t} &= f_t - \bar{f}. \end{aligned}$$

Considerable algebra shows that at the first stage of GMM $T^{-1}V_b^\diamond \xrightarrow{d} \sigma_f^2 Z^2(X'\Sigma_R X)/(X'X)^2$. Hence $t \xrightarrow{d} Z/\sqrt{\sigma_f^2 Z^2(X'\Sigma_R X)/(X'X)^2}$ or $(X'\mu_R)/[\sigma_f^2 Z^2(X'\Sigma_R X)]^{\frac{1}{2}}$. We also have

$$R_\diamond^2 = 1 - \frac{(\bar{R}^e - d_T\hat{b}^\diamond)'(\bar{R}^e - d_T\hat{b}^\diamond)}{\left(\bar{R}^e - \iota\iota'\bar{R}^e/n\right)'\left(\bar{R}^e - \iota\iota'\bar{R}^e/n\right)} = 1 - \frac{\bar{R}^{e\prime} M_d \bar{R}^e}{\bar{R}^{e\prime} M_\iota \bar{R}^e}$$

where $M_d = I - d_T(d_T'd_T)^{-1}d_T'$ and $M_\iota = I_n - \iota\iota'/n$. So the $R^2$ is

$$R_\diamond^2 \xrightarrow{d} 1 - \frac{\mu_R' M \mu_R}{\mu_R' M_\iota \mu_R}$$

where $M = I_n - X(X'X)^{-1}X'$.

At the second stage of GMM the weighting matrix is $W_T^\diamond = (P_T S_T^\diamond P_T')^{-1}$ where $P_T = \begin{pmatrix} I_n & \bar{R}^e(\hat{b}^\diamond)' \end{pmatrix}$. Considerable algebra shows that $TW_T^\diamond \xrightarrow{d} W = \Sigma_R^{-1}/(\sigma_f^2 Z^2)$. We have $T^{-\frac{1}{2}}\hat{b}^\diamond = T^{\frac{1}{2}} d_T'(TW_T)\bar{R}^e/(T^{\frac{1}{2}} d_T'(TW_T)T^{\frac{1}{2}} d_T)$. It follows that $T^{-\frac{1}{2}}\hat{b}^\diamond \xrightarrow{d} \tilde{Z} = (X'W\mu_R)/(X'WX) = (X'\Sigma_R^{-1}\mu_R)/(X'\Sigma_R^{-1}X)$. The $t$-statistic for $\hat{b}^\diamond$ is $t = \hat{b}^\diamond/\sqrt{V_b^\diamond}$ where $V_b^\diamond$ is again the first element on the diagonal of $V_\theta^\diamond$, and $V_\theta^\diamond$ is given by (38). In this case, however, the matrix $a_T^\diamond$ depends on the weighting matrix and takes a form such that $T^{-1}V_b^\diamond \xrightarrow{d} \sigma_f^2 Z^2 (X'W\Sigma_R WX)/(X'WX)^2$ or $\sigma_f^2 Z^2/(X'\Sigma_R^{-1}X)$. Hence $t \xrightarrow{d} \tilde{Z}/[\sigma_f^2 Z^2/(X'\Sigma_R^{-1}X)]^{\frac{1}{2}}$ or $(X'\Sigma_R^{-1}\mu_R)/[\sigma_f^2 Z^2 (X'\Sigma_R^{-1}X)]^{\frac{1}{2}}$. We also have

$$R_\diamond^2 = 1 - \frac{\bar{R}^{e\prime}\tilde{M}_d'\tilde{M}_d\bar{R}^e}{\bar{R}^{e\prime}M_\iota\bar{R}^e}$$

where $\tilde{M}_d = I - d_T(d_T'W_T d_T)^{-1}d_T'W_T$. So the $R^2$ is

$$R_\diamond^2 \xrightarrow{d} 1 - \frac{\mu_R'\tilde{M}'\tilde{M}\mu_R}{\mu_R' M_\iota \mu_R}$$

where $\tilde{M} = I_n - X(X'WX)^{-1}X'W = I_n - X(X'\Sigma_R^{-1}X)^{-1}X'\Sigma_R^{-1}$. The test statistic for the over-identifying restrictions is $J = T(\bar{R}^e - d_T\hat{b})'W_T(\bar{R}^e - d_T\hat{b}) = T\bar{R}^{e\prime}\tilde{M}_d'W_T\tilde{M}_d\bar{R}^e$. Hence $J \xrightarrow{d} \mu_R'\tilde{M}'W\tilde{M}\mu_R = \mu_R'\tilde{M}'\Sigma_u^{-1}\tilde{M}\mu_R/(\sigma_f^2 Z^2)$.■

## 7.6 Proof of Theorem 6

As in the proof to Theorem 3, $\hat{b}^* \xrightarrow{a.s.} b_s^* = (D_0'W_0^*D_0)^{-1} D_0'W_0^*E(R^e)$. When the risk factor is panel spurious $D_0 = c\iota + E(R^e)E(f)$. Hence $b_s^* = (1-\omega_0^*)/E(f)$ where

$$\omega_0^* = \frac{D_0'W_0^*\iota}{D_0'W_0^*D_0}c.$$

The predicted expected returns are $D_T\hat{b}^* \xrightarrow{a.s.} D_0 b_s^*$ and

$$D_0 b_s^* = [E(R^e)E(f) + c\iota]\,\frac{1-\omega_0^*}{E(f)} = \left[E(R^e) + \frac{c}{E(f)}\iota\right](1-\omega_0^*)$$

Notice that $\lim_{c\to 0}^* \omega_0^* = 0$ so that $\lim_{c\to 0} b_s^* = 1/E(f)$ and $\lim_{c\to 0} D_0 b_s^* = E(R^e)$.

Similarly, $\hat{b}^\diamond \xrightarrow{a.s.} b_s^\diamond = (d_0'W_0^\diamond d_0)^{-1} d_0'W_0^\diamond E(R^e)$. When the risk factor is panel spurious $d_0 = c\iota$ hence $b_s^\diamond = \omega_0^\diamond/c$ where

$$\omega_0^\diamond = \frac{\iota'W_0^\diamond E(R^e)}{\iota'W_0^\diamond\iota}.$$

The predicted expected returns are $d_T\hat{b}^\diamond \xrightarrow{a.s.} d_0 b_s^\diamond = \iota\omega_0^\diamond$.■

## 7.7 Inference about the Price of Risk

In what follows it is useful to define a vector $\varsigma_f$ which contains the unique elements of $\Sigma_f$. There are $K = k(k+1)/2$ unique elements of $\Sigma_f$, so $\varsigma_f$ is the $K \times 1$ vector:

$$\varsigma_f = (\overbrace{\sigma_{f,11} \quad \sigma_{f,12} \quad \cdots \quad \sigma_{f,1k}}^{k \text{ elements}} \overbrace{\sigma_{f,22} \quad \sigma_{f,23} \quad \cdots \quad \sigma_{f,2k}}^{k-1 \text{ elements}} \overbrace{\sigma_{f,33} \quad \cdots \quad \sigma_{f,kk}}^{\cdots} )'. \tag{39}$$

Below it will be useful to have a pair of mappings $i(\kappa)$ and $j(\kappa)$ which map the row index, $\kappa$, of an element of $\varsigma_f$ to the corresponding row and columns indices of $\Sigma_f$. These functions can be represented using vectors:

$$\begin{aligned}
\kappa &= (\overbrace{1 \quad 2 \quad \cdots \quad k}^{k \text{ elements}} \overbrace{k+1 \quad k+2 \quad \cdots \quad 2k-1}^{k-1 \text{ elements}} \overbrace{2k \quad \cdots \quad K}^{\cdots} )' \\
i(\kappa) &= (\overbrace{1 \quad 1 \quad \cdots \quad 1}^{k \text{ elements}} \overbrace{2 \quad 2 \quad \cdots \quad 2}^{k-1 \text{ elements}} \overbrace{3 \quad \cdots \quad k}^{\cdots} )' \\
j(\kappa) &= (\overbrace{1 \quad 2 \quad \cdots \quad k}^{k \text{ elements}} \overbrace{2 \quad 3 \quad \cdots \quad k}^{k-1 \text{ elements}} \overbrace{3 \quad \cdots \quad k}^{\cdots} )'
\end{aligned}$$

### 7.7.1 The $a$-Normalization

To estimate $\lambda_f^* = \Sigma_f b^*/(1-\mu' b^*)$ add the moment restrictions (16) and (17). I use a GMM estimator for $b^*$, $\mu$ and $\varsigma_f$ that sets $\tilde{a}_T^* \tilde{g}_T^* = 0$ where

$$\tilde{a}_T^* = \begin{pmatrix} a_T^* & 0 \\ 0 & I_{k+K} \end{pmatrix}, \qquad \tilde{g}_T^* = \begin{pmatrix} g_T^* \\ \hat{g}_T^* \end{pmatrix} \tag{40}$$

and

$$\hat{g}_T^*(\mu, \varsigma_f) = \begin{pmatrix} \bar{f} - \mu \\ (\bar{f}_1 - \mu_1)^2 + s_{f,11} - \sigma_{f,11} \\ (\bar{f}_1 - \mu_1)(\bar{f}_2 - \mu_2) + s_{f,12} - \sigma_{f,12} \\ \vdots \\ (\bar{f}_1 - \mu_1)(\bar{f}_k - \mu_k) + s_{f,1k} - \sigma_{f,1k} \\ (\bar{f}_2 - \mu_2)^2 + s_{f,22} - \sigma_{f,22} \\ (\bar{f}_2 - \mu_2)(\bar{f}_3 - \mu_3) + s_{f,23} - \sigma_{f,23} \\ \vdots \\ (\bar{f}_k - \mu_k)^2 + s_{f,kk} - \sigma_{f,kk} \end{pmatrix}.$$

where $s_{f,ij}$ is the $ij$ element of $S_f = T^{-1} \sum_{t=1}^{T} (f_t - \bar{f})(f_t - \bar{f})'$. The GMM estimators are $\hat{b}^*$, as before, along with $\hat{\mu} = \bar{f}$ and $\hat{\sigma}_{f,ij} = S_{f,ij}$.

For notational convenience let $\tilde{\theta}^* = (\; b^{*\prime} \quad \mu' \quad \varsigma_f' \;)'$ and define

$$\tilde{u}_t^*(\tilde{\theta}^*) = \begin{pmatrix} u_t^*(b^*) \\ \hat{u}_t^*(\mu, \varsigma_f) \end{pmatrix}$$

where

$$\hat{u}_t^*(\mu,\varsigma_f)=\begin{pmatrix} f_t-\mu \\ (f_{1t}-\mu_1)^2-\sigma_{f,11} \\ (f_{1t}-\mu_1)(f_{2t}-\mu_2)-\sigma_{f,12} \\ \vdots \\ (f_{1t}-\mu_1)(f_{kt}-\mu_k)-\sigma_{f,1k} \\ (f_{2t}-\mu_2)^2-\sigma_{f,22} \\ (f_{2t}-\mu_2)(f_{3t}-\mu_3)-\sigma_{f,23} \\ \vdots \\ (f_{kt}-\mu_k)^2-\sigma_{f,kk} \end{pmatrix}.$$

and let where $\sigma_{f,ij}$ is the $ij$ element of $\Sigma_f$.

Let

$$\tilde{a}_0^*=\begin{pmatrix} a_0^* & 0 \\ 0 & I_{k+K} \end{pmatrix}$$

$$\tilde{\delta}_0^*=E\left[\frac{\partial\tilde{u}_t^*(\tilde{\theta}_0^*)}{\partial\tilde{\theta}_0^*}\right]=\begin{pmatrix} \delta_0^* & 0 \\ 0 & -I_{k+K} \end{pmatrix}$$

and

$$\tilde{S}_0^*=E\left[\sum_{j=-\infty}^{+\infty}\tilde{u}_t^*(\tilde{\theta}_0^*)\tilde{u}_{t-j}^*(\tilde{\theta}_0^*)'\right].$$

Letting $\tilde{\theta}_T^*=\begin{pmatrix}\hat{b}^{*\prime} & \hat{\mu}' & \hat{\varsigma}_f'\end{pmatrix}'$, $\sqrt{T}(\tilde{\theta}_T^*-\tilde{\theta}_0^*)\xrightarrow{d}N(0,V_{\tilde{\theta}}^*)$ with $V_{\tilde{\theta}}^*=(\tilde{a}_0^*\tilde{\delta}_0^*)^{-1}\tilde{a}_0^*\tilde{S}_0^*\tilde{a}_0^{*\prime}(\tilde{\delta}_0^{*\prime}\tilde{a}_0^*)^{-1}$.

Now $\lambda_f^*=\Sigma_f b^*/(1-\mu' b^*)$. Hence

$$q^*(\tilde{\theta}^*)=\frac{d\lambda_f^*(\tilde{\theta}^*)}{d\tilde{\theta}^*}=\begin{pmatrix}\Sigma_f+\lambda_f\mu' & \lambda_f b^{*\prime} & \Psi^*\end{pmatrix}/(1-\mu' b^*)$$

where $\Psi_{i(\kappa),\kappa}^*=b_{j(\kappa)}^*$, $\Psi_{j(\kappa),\kappa}^*=b_{i(\kappa)}^*$, $\kappa=1,\ldots,K$ and all other entries in $\Psi^*$ are zero. By the delta method it follows that $\sqrt{T}(\hat{\lambda}_f^*-\lambda_{f0})\xrightarrow{d}N(0,q_0^*V_{\tilde{\theta}}^*q_0^{*\prime})$ where $q_0^*=q^*(\tilde{\theta}_0^*)$.

### 7.7.2 The $\xi$-Normalization

To estimate $\lambda_f^\diamond=\Sigma_f b^\diamond$ add the moment restrictions (17). I use a GMM estimator for $b^\diamond$, $\mu$ and $\varsigma_f$ that sets $\tilde{a}_T^\diamond\tilde{g}_T^\diamond=0$ where

$$\tilde{a}_T^\diamond=\begin{pmatrix} a_T^\diamond & 0 \\ 0 & I_K \end{pmatrix},\qquad \tilde{g}_T^\diamond=\begin{pmatrix} g_T^\diamond \\ \hat{g}_T^\diamond \end{pmatrix}$$

and $\hat{g}_T^\diamond(\mu,\varsigma_f)$ is the last $K$ rows of $\hat{g}_T^\diamond(\mu,\varsigma_f)$. The GMM estimators are $\hat{b}^\diamond$, $\hat{\mu}=\bar{f}$, as before, along with and $\hat{\sigma}_{f,ij}=\mathrm{S}_{f,ij}$.

For notational convenience let $\tilde{\theta}^{\diamond} = ( \begin{array}{cc} \theta^{\diamond\prime} & \varsigma_f^{\prime} \end{array} )^{\prime}$ and define

$$\tilde{u}_t^{\diamond}(\tilde{\theta}^{\diamond}) = \left( \begin{array}{c} u_t^{\diamond}(\theta^{\diamond}) \\ \hat{u}_t^{\diamond}(\mu, \varsigma_f) \end{array} \right)$$

where $\hat{u}_t^{\diamond}(\mu, \varsigma_f)$ is the last $K$ rows of $\hat{u}_t^*(\mu, \varsigma_f)$.

Let

$$\tilde{a}_0^{\diamond} = \left( \begin{array}{cc} a_0^{\diamond} & 0 \\ 0 & I_K \end{array} \right)$$

$$\tilde{\delta}_0^{\diamond} = E\left[\frac{\partial \tilde{u}_t^{\diamond}(\tilde{\theta}_0^{\diamond})}{\partial \tilde{\theta}_0^{*}}\right] = \left( \begin{array}{cc} \delta_0^{\diamond} & 0 \\ 0 & -I_K \end{array} \right)$$

and

$$\tilde{S}_0^{\diamond} = E\left[\sum_{j=-\infty}^{+\infty} \tilde{u}_t^{\diamond}(\tilde{\theta}_0^{\diamond}) \tilde{u}_{t-j}^{\diamond}(\tilde{\theta}_0^{\diamond})^{\prime}\right].$$

Letting $\tilde{\theta}_T^{\diamond} = ( \begin{array}{cc} \hat{\theta}^{\diamond\prime} & \hat{\varsigma}_f^{\prime} \end{array} )^{\prime}$, $\sqrt{T}(\tilde{\theta}_T^{\diamond} - \tilde{\theta}_0^{\diamond}) \xrightarrow{d} N(0, V_{\tilde{\theta}}^{\diamond})$ with $V_{\tilde{\theta}}^{\diamond} = (\tilde{a}_0^{\diamond}\tilde{\delta}_0^{\diamond})^{-1}\tilde{a}_0^{\diamond}\tilde{S}_0^{\diamond}\tilde{a}_0^{*\prime}(\tilde{\delta}_0^{\diamond\prime}\tilde{a}_0^{\diamond})^{-1}$.

Now $\lambda_f^{\diamond} = \Sigma_f \hat{b}^{\diamond}$. Hence

$$q^{\diamond}(\tilde{\theta}^{\diamond}) = \frac{d\hat{\lambda}_f^{\diamond}(\tilde{\theta}^{\diamond})}{d\tilde{\theta}^{\diamond}} = \left( \begin{array}{ccc} \Sigma_f & 0_{k \times k} & \Psi^{\diamond} \end{array} \right)$$

where $\Psi_{i(\kappa),\kappa}^{\diamond} = b_{j(\kappa)}^{\diamond}$, $\Psi_{j(\kappa),\kappa}^{\diamond} = b_{i(\kappa)}^{\diamond}$, $\kappa = 1, \ldots, K$ and all other entries in $\Psi^{\diamond}$ are zero. By the delta method it follows that $\sqrt{T}(\hat{\lambda}_f^{\diamond} - \lambda_{f0}) \xrightarrow{d} N(0, q_0^{\diamond} V_{\tilde{\theta}}^{\diamond} q_0^{\diamond\prime})$ where $q_0^{\diamond} = q^{\diamond}(\tilde{\theta}_0^{\diamond})$.

## 7.8 Estimating Long-Run Covariance Matrices

### 7.8.1 The $a$-Normalization

As stated in section 4, I define $S_T^* = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^* \hat{u}_t^{*\prime}$ when estimating the standard errors of $\hat{b}^*$ and testing the over-identifying restrictions of the model. This is a consistent estimate of $S_0^*$ when the model is true because $E[u_t^*(b_0^*) u_{t-j}^*(b_0^*)^{\prime}] = 0$ for $j \neq 0$.

When conducting inference about the price of risk we need an estimate of $\tilde{S}_0^*$. Since $\hat{u}_t^*(\mu_0, \varsigma_{f0})$ is not necessarily orthogonal to lagged information the simple covariance matrix $\frac{1}{T}\sum_{t=1}^{T} \tilde{u}_t^*(\tilde{\theta}_T^{\diamond})\tilde{u}_t^*(\tilde{\theta}_T^{\diamond})^{\prime}$ will, in general, be an inconsistent estimator of $\tilde{S}_0^*$. For this reason I use den Haan and Levin's (2000) VARHAC procedure for estimating $\tilde{S}_0^*$. In doing so I impose the restriction that lagged variables do not appear in the equations for $u_t^*$ but allow for lags in the equations for $\hat{u}_t^*$.

### 7.8.2 The $\xi$-Normalization

As stated in section 3, to compute $S_T^{\diamond}$ I use the same VARHAC procedure described above. In doing so I impose the restriction that lagged variables do not appear in the equations for $u_{1t}^{\diamond}$ (the errors corresponding to the asset pricing conditions) but allow for lags in the equations for $u_{2t}^{\diamond *}$ (the errors corresponding to $f_t - \mu$). When conducting inference about the price of risk I continue to use the VARHAC procedure to estimate $\tilde{S}_0^{\diamond}$. In doing so I allow for lags in the equations for $\hat{u}_t^*$.

## 7.9 Proof of Theorem 7

When the model if false, the estimated asymptotic variance-covariance matrix of $\tilde{\theta}_T^*$ is $\hat{V}_{\tilde{\theta}}^* = (\tilde{a}_T^*\tilde{\delta}_T^*)^{-1}\tilde{a}_T^*\tilde{S}_T^*\tilde{a}_T^{*\prime}(\tilde{\delta}_T^{*\prime}\tilde{a}_T^*)^{-1}$ where $\delta_T^* = -D_T$, $\tilde{a}_T^*$ is given by (40) and $\tilde{S}_T^*$ is the VARHAC estimator described above. Consequently

$$\hat{V}_{\tilde{\theta}}^* \xrightarrow{a.s.} V_{\tilde{\theta}s}^* = \begin{pmatrix} (D_0'W_0^*D_0)^{-1}D_0'W_0^* & 0 \\ 0 & I_{k+K} \end{pmatrix} \tilde{V}_s^* \begin{pmatrix} W_0^*D_0(D_0'W_0^*D_0)^{-1} & 0 \\ 0 & I_{k+K} \end{pmatrix}.$$

Here $\tilde{V}_s^* = \text{plim}\,\tilde{S}_T^*$ and does not necessarily correspond to $\tilde{S}_s^* = E\left[\sum_{j=-\infty}^{+\infty}\tilde{u}_t^*(\tilde{\theta}_s^*)\tilde{u}_{t-j}^*(\tilde{\theta}_s^*)'\right]$, where $\tilde{\theta}_s^* = \text{plim}\,\tilde{\theta}_T^*$ because lag restrictions are imposed in computing $\tilde{S}_T^*$ that are not true when the model is false. Importantly, however, $V_{\tilde{\theta}s}^*$ is finite.

The squared $t$ statistics for $\hat{\lambda}_f^*$ are

$$\begin{aligned} t^2 &= \frac{T\,\text{diag}(\hat{\lambda}_f^*\hat{\lambda}_f^{*\prime})(1-\hat{\mu}'\hat{b}^*)^2}{\text{diag}\left[\begin{pmatrix} \hat{\Sigma}_f + \hat{\lambda}_f^*\hat{\mu}' & \hat{\lambda}_f^*\hat{b}^{*\prime} & \hat{\Psi}^* \end{pmatrix}\hat{V}_{\tilde{\theta}}^*\begin{pmatrix} \hat{\Sigma}_f + \hat{\lambda}_f^*\hat{\mu}' & \hat{\lambda}_f^*\hat{b}^{*\prime} & \hat{\Psi}^* \end{pmatrix}'\right]} \\ &= \frac{T\,\text{diag}(\hat{\Sigma}_f\hat{b}^*\hat{b}^{*\prime}\hat{\Sigma}_f)(1-\hat{\mu}'\hat{b}^*)^2}{\text{diag}(X\hat{V}_{\tilde{\theta}}^*X')}. \end{aligned}$$

where $X = \begin{pmatrix} (1-\hat{\mu}'\hat{b}^*)\hat{\Sigma}_f + \hat{\Sigma}_f\hat{b}^*\hat{\mu}' & \hat{\Sigma}_f\hat{b}^*\hat{b}^{*\prime} & (1-\hat{\mu}'\hat{b}^*)\hat{\Psi}^* \end{pmatrix}$ and the division is element-by-element. Notice that $\hat{\Sigma}_f\hat{b}^* \xrightarrow{a.s.} \Sigma_{f0}b_s^*$ while $X \xrightarrow{a.s.} \Sigma_{f0}b_s^*\begin{pmatrix} \mu' & b_s^{*\prime} & 0 \end{pmatrix}$. Hence $t^2$ has the same asymptotic distribution as

$$\tilde{t}^2 = T\frac{(1-\hat{\mu}'\hat{b}^*)^2\iota}{\text{diag}\left[\begin{pmatrix} \mu' & b_s^{*\prime} & 0 \end{pmatrix}V_{\tilde{\theta}s}^*\begin{pmatrix} \mu' & b_s^{*\prime} & 0 \end{pmatrix}'\right]},$$

where $\iota$ is a $k \times 1$ vector of ones. If the lag restrictions imposed in computing $\tilde{S}_T^*$ are valid then each element in $t^2$ converges in distribution to a $\chi^2$ with 1 degree of freedom. From this we can see that a researcher will conclude that $\hat{\lambda}_f^*$ is significantly different from 0 in roughly $\alpha$ percent of repeated samples if he uses an $\alpha$ percent critical value in this test. But if the lag restrictions are not valid, the rate at which the researcher will reject $\lambda = 0$ could be higher or lower than $\alpha$ percent.■

# 8 Appendix B: Data Construction

## 8.1 Fama-French Portfolios

Each Fama and French (1993) portfolio represents the intersection of one of 5 groups of stocks sorted according to their market capitalization with one of 5 groups of stocks sorted according to their book equity to market capitalization ratio. The returns are equally weighted. I obtained raw monthly returns from Kenneth French's website http://mba.tuck.dartmouth.edu/ pages/ faculty/ ken.french/ data_library.html. To obtain quarterly returns I compounded monthly returns within each quarter. To obtain quarterly excess returns I subtract the quarterly risk free rate defined as the compounded monthly risk free rate from Fama/French Research Data Factor file. Real excess returns are defined by dividing the nominal excess return by one plus the inflation rate, which I defined below.

To compute real consumption of nondurables and services I proceeded as follows. Let $C_t^N$ be the consumption of nondurables and $C_t^S$ be the consumption of services in nominal dollars, and let $c_t^N$ and $c_t^S$ be the corresponding series in constant chained dollars, as published by the Bureau of Economic Analysis. To obtain nominal consumption of nondurables and services I simply set $C_t = C_t^N + C_t^S$. However, because real chained series are not summable, I proceed as follows to create real consumption of nondurables and services, which I denote $c_t$. First define $s_t = (C_t^N/C_t + C_{t-1}^N/C_{t-1})$, $g_t^N = c_t^N/c_{t-1}^N - 1$ and $g_t^S = c_t^S/c_{t-1}^S - 1$. Then define the growth rate of $c_t$ as $g_t = s_t g_t^N + (1 - s_t) g_t^S$. Notice that a real levels series can then be generated by forward and backward induction relative to a base period. I convert the real levels series into per capita terms by dividing by the quarterly population series published in the National Income and Product Accounts by the BEA.[9] I construct an inflation series using a similar method. Letting $\pi_t^N$ and $\pi_t^S$ be the inflation rates for nondurables and services, I let the combined inflation rate be $\pi_t = s_t \pi_t^N + (1 - s_t)\pi_t^S$.

I assume that households derive utility in quarter $t + 1$ from the stock of durables at the end of quarter $t$. To compute the real quarterly stock of durable goods I proceeded

[9] I pass the NIPA population series through the Census X11 seasonal adjustment procedure because the NIPA series displays noticeable seasonal fluctuations.

as follows. The Bureau of Economic Analysis publishes end-of-year real stocks of durables goods. Let $k_t$ denote the real stock of durables at the end of some year, and let $k_{t+4}$ be the same stock a year (four quarters) later. We observe quarterly real purchases of consumer durables, which I denote $c_t^D$. I assume that within each year the model

$$k_{t+1} = c_{t+1}^D + (1-\delta)k_t \tag{41}$$

holds, with $\delta$ allowed to vary by year. I solve for the value of $\delta$ such that the beginning and end-of-year stocks are rationalized by purchases series. This is the $\delta$ such that

$$k_{t+4} = c_{t+4}^D + (1-\delta)c_{t+3}^D + (1-\delta)^2 c_{t+2}^D + (1-\delta)^3 c_{t+1}^D + (1-\delta)^4 k_t. \tag{42}$$

Once I identify the value of $\delta$ that applies within a year using (42), I use (41) to calculates the within year stocks. I convert the real stocks to per capita terms by dividing by the same population series used for the consumption series.

The remaining series are taken from the Fama/French Research Data Factor file. I defined the monthly market return as the sum of the market premium series (*RM-Rf*) and the risk free rate series (*Rf*). I convert this to a quarterly return by compounding the monthly series geometrically within each quarter. Denoting the resulting series, $R_t^M$, I convert it to a real return as follows: $r_t^M = (R_t^M - \pi_t)/(1+\pi_t)$.

To create real quarterly versions of the Fama-French factors (*RM-Rf*, *SMB* and *HML*) I proceed as follows. To get quarterly excess returns I compound the monthly series geometrically within each quarter. I convert them to real excess returns by dividing the resulting series by $1 + \pi_t$.

## 8.2 Lustig-Verdelhan Foreign Exchange Portfolios

I take all of the data directly from LV's database provided at http://www.econ.ucla.edu/hlustig/. The construction of the eight portfolio returns is described fully in LV. The risk factors are also defined exactly as in their paper. I have additional results—which are qualitatively similar to the ones presented here—based on the same portfolio returns, but using alternative measures of the risk factors derived from the same sources as for the quarterly data. For the annual series consumption of nondurables and services is constructed using the annual NIPA data, using the same approach as for the quarterly data. The annual stocks of durables are available directly from the NIPA accounts. The annual Fama/French risk factors and the risk free rate are also available directly from the Fama/French Research Data Factor file.

# 9 Appendix C: Small-Sample P-Values for Covariance and Beta Tests

To obtain small sample p-values relevant to the LV data, I ran 5000 Monte Carlo simulations, each with a sample size of $T = 50$. The simulated $R_t^e$ process was distributed as an iid $N(\mu_R, \Sigma_R)$, where $\mu_R$ and $\Sigma_R$ were set equal to their sample equivalents in the LV data. Since the null being tested is whether the covariance of the factor is zero with all the returns, I simulated a risk factor $f_t$ which, in the case of consumption growth, was set to an AR(2) process with the same mean, variance and autoregressive coefficients as US consumption growth in the LV data, and was made independent of $R_t^e$. Similar simulations were run to generate factors that mimicked the univariate representations of durables growth [also specified as an AR(2)], and the market return (specified as iid). In each sample I computed the Wald statistics for the hypotheses that $\text{cov}(R_{it}^e, f_t) = 0$ for all $i$, $\text{cov}(R_{it}^e, f_t) = c$ for all $i$, $\beta_i = 0$ for all $i$, and $\beta_i = \beta$ for all $i$ and thereby constructed empirical distributions for the test statistics across the 5000 simulations. I used this distribution to back out the small sample p-values presented in Tables 5 and 6.

# 10 Appendix D: Robustness

#### 10.0.1 Covariance Matrix Estimation

In estimating the models I used the VARHAC procedure of den Haan and Levin (2000) to compute the standard errors and to compute the long-run covariance matrices that are used in constructing the second stage GMM estimates. This procedure is mainly important when the moments of the risk factors are estimated because the error terms for these moment conditions display some degree of serial correlation, especially for the consumption variables. To determine whether the VARHAC procedure plays a significant role in driving the results, I redid the analysis using a simple HAC procedure that allows for no lags.

For the most part, the results—presented in Tables 1a, 2a, 3a and 4a—are very similar to the benchmark results. For both sets of data, if anything, using HAC standard errors enhances the contrast between the results for the two normalizations because there are more rejections of the over-identifying restrictions for the $\xi$-normalization. For the LV data using HAC procedure tends to reduce the standard errors of the estimates of $\hat{\lambda}_f$, while, at the same time, worsening the fit of the model.

### 10.0.2 Different Weighting Matrices

As described earlier, when estimating the model using the $\xi$-normalization, I set the first stage weighting matrix for the asset pricing conditions to $W_T^{\diamond} = I_n$. In the second stage, I let $W_T^{\diamond} = (P_T S_T^{\diamond} P_T')^{-1}$ where $P_T = \begin{pmatrix} I_n & \bar{R}^e(\hat{b}^{\diamond})' \end{pmatrix}$ and $S_T^{\diamond}$ is a consistent estimator of the long-run covariance matrix of the GMM errors. The second entry in $P_T$ is the derivative of the pricing errors, $\bar{R}^e - \left(D_T - \bar{R}^e\mu'\right) b^{\diamond}$, with respect to $\mu$. This derivative also appears in the definition of $\delta_T^{\diamond}$, given by equation (22), which is used in computing standard errors and the test of the over-identifying restrictions. Under the null that the model is true, of course, $\text{plim}\, \bar{R}^e = \text{plim}\, d_T \hat{b}^{\diamond}\hat{b}^{\diamond\prime}$. This observation leads to two alternative schemes for estimating the $\xi$-normalization.

In the first scheme $\bar{R}^e(\hat{b}^{\diamond})'$ is replaced by $d_T\hat{b}^{\diamond}\hat{b}^{\diamond\prime}$ in the definition of both $P_T$ and $\delta_T^{\diamond}$ but the procedures for estimating the model are otherwise unchanged.

In the second scheme, suggested by Cochrane (2005), $\bar{R}^e(\hat{b}^{\diamond})'$ is replaced by $d_T\hat{b}^{\diamond}\hat{b}^{\diamond\prime}$ in the definition of $\delta_T^{\diamond}$. However, Cochrane defines the second stage weighting matrix as $W_T^{\diamond} = \left(\tilde{P} S_T^{\diamond} \tilde{P}'\right)^{-1}$ where $\tilde{P} = \begin{pmatrix} I_n & 0_{n\times k} \end{pmatrix}$.

I re-estimated the model using both of these schemes. For the most part the results are very similar to the base case. Results for the first scheme are presented in Tables 2b and 4b. Results for the second scheme are presented in Tables 2c and 4c. The sharp contrast between the fit of the two normalizations is preserved. For the LV data both procedures notably reduce the standard errors of $\hat{\lambda}_f$, while worsening the fit of the model. For the FF25 data the first scheme appears to systematically increase the magnitude of the estimated $b^{\diamond}$ coefficients, rendering some of them statistically significant for the consumption-based models. However, the fit of the model remains very poor in each of these cases.

### 10.0.3 Models with a Common Pricing Error

Parker and Julliard (2005) use the $\xi$-normalization but modify the moment conditions by adding a common pricing error, $\gamma$, to each of the asset pricing conditions. They use the moment conditions

$$E(f_t) = \mu \qquad E\left\{R_t^e[1-(f_t-\mu)'b^\diamond]-\gamma\right\} = 0, \tag{43}$$

in place of (9). In theory, of course, $\gamma = 0$ if the original model, (1), is true. Parker and Julliard include the additional parameter to "separately evaluate the ability of the model to explain the equity premium and the cross section of expected stock returns." As they suggest, $\hat{\gamma}$ "measures the extent to which the model underpredicts the excess returns of all ... portfolios by the same amount."

Parker and Julliard's modification of the $\xi$-normalization is akin to the use of a constant in the cross-sectional regression of the two-pass procedure. In fact, the two procedures are numerically equivalent when $W_T^\diamond = I_n$ [see Burnside (2007a)].

I re-estimated the model using the moment conditions (43), while leaving the rest of the benchmark procedure unchanged. There are some noticeable differences in the results, which are presented in Tables 2d and 4d. First, for both sets of data, not surprisingly, the fit of the models generally improves. This is guaranteed to happen in the first stage of GMM because the benchmark estimates are equivalent to a cross-sectional regression of $\bar{R}^e$ on $d_T$ without a constant, while the alternative approach is equivalent to adding a constant to the regression. Nonetheless, the fit of the models, with one exception, remains worse than the fit implied by estimates of the $a$-normalization (compare Table 2d to Table 1, and Table 4d to Table 3). The one exception is the single factor model with durables growth as the factor for the FF25 data.

For the LV data, another distinction between the benchmark results in Table 2 and the alternative results in Table 2d is that in the latter case the standard errors for $\hat{b}^\diamond$ and $\hat{\lambda}_f$ are larger, at least for the consumption based models. Also, none of the consumption-based models can be rejected on the basis of the test of the over-identifying restrictions. This is not, however, a virtue of the models. It reflects the fact that the relative precision in the estimates in Table 2 stems from the model having to match the average excess return across portfolios. Once a common pricing error is introduced in the model, the model parameters are effectively unidentified.

In contrast, for the FF25 data, the standard errors of $\hat{b}^\diamond$ and $\hat{\lambda}_f$ are generally smaller for the alternative results (Table 4d) than for the benchmark case (Table 4), but only for consumption based models. Nonetheless, none of the consumption factors are statistically significant—in terms of either $\hat{b}^\diamond$ or $\hat{\lambda}_f$—in any of the models. For the single factor model with the market premium, and for the multi-factor model with the Fama-French factors, there are sign reversals for the coefficients on the market premium and the *SMB* factor. But the fit of the models remains roughly the same. Finally, a common feature to all the models is that the constant pricing error, $\hat{\gamma}$, is large and statistically significant in every case. The estimates range in magnitude from 1.85 percent to 3.83 percent and these are quarterly rates. This suggests that there are either severe small sample problems in estimation, with the maintained hypothesis being that the models are true, or it indicates that the models are misspecified. A third possibility, that the returns to treasury bills mismeasure the risk free rate, is less plausible, because of the magnitude of the coefficients.

### 10.0.4 Using All Information to Estimate $\mu$

Each of the variants of the $\xi$-normalization that I have described above sets the GMM estimator up in such a way that $\hat{\mu} = \bar{f}$. Consider the benchmark case where the GMM estimator sets $a_T^\diamond g_T^\diamond = 0$ where $g_T^\diamond = \begin{pmatrix} g_{1T}^{\diamond\prime} & g_{2T}^{\diamond\prime} \end{pmatrix}'$, $g_{1T}^\diamond = \bar{R}^e - \left(D_T - \bar{R}^e\mu'\right) b^\diamond$, $g_{2T}^\diamond = \bar{f} - \mu$ and $a_T^\diamond$ is given by (19). A more traditional approach to GMM might, instead, define

$$a_T^\diamond = \begin{pmatrix} (D_T - \bar{R}\mu') & -\bar{R}^e b^{\diamond\prime} \\ 0 & I_k \end{pmatrix} W_T^\diamond \tag{44}$$

where $W_T^\diamond$ would now be an $(n+k)\times(n+k)$ weighting matrix. With the $a_T^\diamond$ given by (44), the equation $a_T^\diamond g_T^\diamond = 0$ is the first order condition corresponding to $\min_{b^\diamond,\mu} g_T^{\diamond\prime} W_T^\diamond g_T^\diamond$. It is clear that in this setup, $\mu$ is free to help match not only $E(f_t - \mu) = 0$ but also the asset pricing equations. Under the null, this is asymptotically more efficient than the other approaches because it uses information about $\mu$ that lies in the asset pricing restrictions.

Yogo (2006) uses this approach. In the first stage of GMM he sets

$$W_T^\diamond = \begin{pmatrix} \kappa I_n & 0 \\ 0 & \mathrm{S}_f^{-1} \end{pmatrix}$$

and $\kappa = \det(\mathrm{S}_{R^e})^{-1/n}$. Here $\mathrm{S}_{R^e} = \frac{1}{T}\sum_{t=1}^{T}(R_t^e - \bar{R}^e)(R_t^e - \bar{R}^e)'$. As $\kappa \to 0$, $\hat{\mu} \to \bar{f}$, whereas, as $\kappa \to \infty$, $\hat{\mu}$ is determined solely by the asset pricing conditions. To give some sense of the magnitudes involved, for the single-factor consumption model and FF25 returns, $\mathrm{S}_f^{-1} = 38447$ and $\kappa = 1097$. In the second stage of GMM, the inverse of a consistent estimate of $S_0^\diamond$ is used as the weighting matrix.

Not surprisingly, using this method makes the models perform better overall because the estimated means of the factors move substantially to fit the asset pricing conditions. Table

2e presents results that should be compared to Table 2 for the LV data. For the single-factor consumption model, the estimate of $\hat{\mu}$ moves almost two standard errors away from $\bar{f}$. The statistical significance of $\hat{b}^{\diamond}$ and $\hat{\lambda}_f$ sharpens slightly at the second stage of GMM and the fit of the model improves, but it is still rejected at below the 5 percent level. For durables growth a similar pattern is seen, and the model is no longer rejected at the 5 percent level. For the market premium, $\hat{\mu}$ moves more than two standard deviations from the mean, but in opposite directions in the two stages of GMM. But the model is still rejected. For the multi-factor models there are significant changes in the estimates of $b^{\diamond}$, $\lambda_f$ and $\mu$, but generally speaking it remains true that consumption-based models with statistically significant values of $b^{\diamond}$ and $\lambda_f$ fit the data poorly, and that point estimates are very sensitive to the choice of weighting matrix.

Table 4e presents results for the FF25 data. The fit of most of the models improves dramatically, compared to the benchmark results in Table 4. Notice, however, that this involves enormous changes in the means of at least one factor in each model. For the basic consumption growth model, $\hat{\mu}$ moves down by about 5 standard errors. For the basic durables model $\hat{\mu}$ moves up by between 3 and 5 standard errors in the two stages of GMM. In one case, the market premium's mean shifts down by almost 7 standard errors. Not surprisingly, in some cases, these shifts in $\hat{\mu}$ go along with big changes in $\hat{b}$ and $\hat{\lambda}_f$. Perhaps the most significant improvement in model performance is for Yogo's multi-factor model. In Table 4, the only significant parameter for this model is the factor risk premium, $\lambda_f$, associated with the market return, *Rm*. In Table 4e, however, the $b$ coefficient for durables growth is significant, and the factor risk premium for durables is also significant at the second stage of GMM. These results are qualitatively similar to Yogo's although the $R^2$ measure of fit is considerably lower here.

Do these results point to a systematic difference between the benchmark procedure and Yogo's approach, or does something else explain Yogo's positive assessment of the model? Yogo uses a shorter sample period (1951:Q1–2001Q4) and also uses Campbell's (2003) timing for consumption growth in which the returns to the equity portfolios in the $n$th quarter are priced using consumption growth between quarter $n$ and $n + 1$. I find, however, that shortening the sample period and switching to Campbell timing has no appreciable effect on the model's fit nor on the statistical significance of $\hat{b}^{\diamond}$ and $\hat{\lambda}_f$. This is true for both the benchmark estimation procedure and Yogo's estimation procedure. Yogo also uses HAC rather than VARHAC covariance matrices. While this affects point estimates, it does not have an appreciable effect on the model's fit nor on the statistical significance of $\hat{b}^{\diamond}$ and $\hat{\lambda}_f$. I do find that second stage GMM estimates are highly sensitive to the choice of $\kappa$ when using Yogo's approach. This sensitivity diminishes if further GMM iterations are performed and HAC standard errors are calculated.

#### 10.0.5 Further Iterations on the Weighting Matrix

As a final check on robustness, I consider iterating on the weighting matrix until approximate convergence in the estimates of $b^*$ and $b^{\diamond}$ is obtained. Results for both normalizations and the LV data are shown in Table 1f/2f. For the $a$-normalization the results are quite similar to the results for two-stage GMM (see Table 1). For the $\xi$-normalization some of the parameter values change substantially, but they are all well within one standard error of the estimates for two-stage GMM (see Table 2). The contrast between the fit of the two normalizations is robust.

Results for both normalizations and the FF25 data are shown in Table 3f/4f. For the $a$-normalization the results are quite similar to the results for two-stage GMM (see Table 3). For the $\xi$-normalization some of the parameter values change substantially, but they are all well within one standard error of the estimates for two-stage GMM (see Table 4). The most noticeable change in the results is that for the $\xi$-normalization every model is rejected at small significance levels after iterating over the weighting matrix, whereas this was not the case for two-stage GMM. The striking contrast between the fit of the two normalizations is robust.

## ADDITIONAL REFERENCE

Campbell, John Y. (2003) "Consumption-Based Asset Pricing," in George M. Constantinides, Milton Harris, and René Stulz, ed.: *Handbook of the Economics of Finance*, Vol. 1B, pp. 801–885. Elsevier: Amsterdam.

TABLE 1a: **GMM Estimates of Linear Factor Models**
**Lustig-Verdelhan Dataset, $a$-Normalization, HAC Standard Errors**

| | First Stage | | | Second Stage | | | |
|---|---|---|---|---|---|---|---|
| | $b^*$ | $\lambda_f$ | $R^2$ | $b^*$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | 48.6 | 4.20 | 0.87 | 48.8 | 4.28 | 0.87 | 5.1 |
| | (10.4) | (3.80) | | (4.3) | (1.33) | | (0.645) |
| **Durables growth** | 20.2 | 2.65 | 0.84 | 24.3 | 5.51 | 0.81 | 4.8 |
| | (3.7) | (1.47) | | (1.7) | (2.32) | | (0.683) |
| ***Rm-Rf*** | 5.5 | 29.13 | 0.46 | 6.5 | 39.05 | 0.45 | 10.2 |
| | (2.3) | (18.05) | | (1.7) | (17.62) | | (0.175) |
| **Consumption Factors** | | | | | | | |
| Consumption growth | 40.1 | 3.64 | 0.87 | -5.5 | 1.86 | 0.81 | 2.7 |
| | (56.0) | (1.76) | | (27.9) | (1.80) | | (0.847) |
| Durables growth | 3.7 | 3.64 | | 26.2 | 4.81 | | |
| | (28.0) | (1.74) | | (13.5) | (2.26) | | |
| **Yogo Factors** | | | | | | | |
| Consumption growth | 9.6 | 2.76 | 0.95 | 2.3 | 3.43 | 0.89 | 1.9 |
| | (30.7) | (1.71) | | (21.9) | (2.47) | | (0.860) |
| Durables growth | 14.2 | 2.91 | | 22.0 | 5.96 | | |
| | (17.1) | (2.34) | | (11.0) | (3.86) | | |
| *Rm* | 2.4 | 33.44 | | 1.1 | 12.09 | | |
| | (1.7) | (23.40) | | (0.9) | (21.59) | | |
| **Fama-French Factors** | | | | | | | |
| *Rm-Rf* | 6.0 | 31.28 | 0.56 | 7.6 | 92.83 | 0.27 | 5.3 |
| | (3.4) | (28.89) | | (2.5) | (93.89) | | (0.384) |
| *SMB* | -5.3 | -12.70 | | -4.5 | -17.52 | | |
| | (5.7) | (21.69) | | (4.2) | (37.41) | | |
| *HML* | 4.1 | 12.25 | | 6.5 | 45.71 | | |
| | (3.9) | (19.33) | | (2.8) | (53.20) | | |

*Note*: Annual data, 1953–2002. The table reports first and second stage GMM estimates of $b^*$, from the SDF $m_t = 1 - f_t' b^*$, obtained using the moment restriction $E(R_t^e m_t) = 0$, where $R_t^e$ is an $8 \times 1$ vector of excess returns of equally-weighted portfolios of short-term foreign-currency denominated money market securities sorted by their interest differential with the US, and $f_t$ is a scalar or vector of risk factors. The factors are real per household consumption (nondurables & services) growth, real per household durable consumption growth, and the following variables from the Fama-French dataset: the real value weighted US stock market excess return over the risk free rate (*Rm-Rf*), the gross return to the same portfolio (*Rm*), and the *SMB* and *HML* portfolio excess returns [see Lustig and Verdelhan (2007)]. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^* / (1 - \bar{f}' \hat{b}^*)$ are also reported (in percent), where $\bar{f}$ and $\mathrm{S}_f$ are the sample mean and covariance matrix of $f_t$. GMM-HAC standard errors are reported in parentheses for $\hat{b}^*$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns, given by $D_T \hat{b}^*$, where $D_T = \frac{1}{T} \sum_{t=1}^{T} R_t^e f_t'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 2a: **GMM Estimates of Linear Factor Models**
**Lustig-Verdelhan Dataset, $\xi$-Normalization, HAC Standard Errors**

| | First Stage | | | | Second Stage | | | |
|---|---|---|---|---|---|---|---|---|
| | $\mu$ | $b^\diamond$ | $\lambda_f$ | $R^2$ | $b^\diamond$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | 0.016 | 45.1 | 0.95 | 0.10 | 76.9 | 1.62 | 0.05 | 17.4 |
| | (0.002) | (71.7) | (1.42) | | (29.8) | (0.73) | | (0.015) |
| **Durables growth** | 0.034 | 20.8 | 0.87 | 0.16 | 59.0 | 2.46 | -0.40 | 17.4 |
| | (0.003) | (28.1) | (1.1) | | (21.3) | (0.8) | | (0.015) |
| ***Rm-Rf*** | 0.070 | 1.8 | 5.97 | 0.02 | 1.4 | 4.51 | 0.02 | 23.3 |
| | (0.026) | (3.5) | (11.37) | | (2.3) | (7.47) | | (0.002) |
| **Consumption Factors** | | | | | | | | |
| Consumption growth | 0.016 | -8.7 | 0.26 | 0.16 | 6.6 | 1.17 | -0.34 | 17.7 |
| | (0.002) | (55.6) | (0.77) | | (39.3) | (0.69) | | (0.007) |
| Durables growth | 0.034 | 23.6 | 0.82 | | 54.5 | 2.40 | | |
| | (0.003) | (38.0) | (1.01) | | (27.5) | (0.80) | | |
| **Yogo Factors** | | | | | | | | |
| Consumption growth | 0.016 | -22.0 | 0.59 | 0.34 | 28.1 | 2.42 | -1.11 | 13.4 |
| | (0.002) | (62.4) | (1.03) | | (45.9) | (1.01) | | (0.020) |
| Durables growth | 0.034 | 45.5 | 1.10 | | 87.4 | 3.86 | | |
| | (0.003) | (49.8) | (1.58) | | (32.0) | (1.14) | | |
| *Rm* | 0.070 | 5.2 | 11.74 | | 4.4 | 8.22 | | |
| | (0.025) | (2.7) | (7.71) | | (2.5) | (7.80) | | |
| **Fama-French Factors** | | | | | | | | |
| *Rm-Rf* | 0.070 | 1.5 | 7.07 | 0.08 | -0.5 | 0.71 | 0.04 | 18.8 |
| | (0.026) | (4.1) | (10.87) | | (3.4) | (9.10) | | (0.002) |
| *SMB* | 0.024 | 1.7 | 4.08 | | 2.4 | 4.03 | | |
| | (0.019) | (4.5) | (6.82) | | (3.8) | (5.82) | | |
| *HML* | 0.057 | -2.8 | -5.91 | | -2.7 | -4.96 | | |
| | (0.020) | (4.9) | (8.03) | | (3.9) | (6.69) | | |

*Note*: Annual data, 1953–2002. The table reports first and second stage GMM estimates of $\mu$ and $b^\diamond$, from the SDF $m_t = 1 - (f_t - \mu)' b^\diamond$, obtained using the moment restrictions $E(R_t^e m_t) = 0$, $E(f_t - \mu) = 0$. Since $\hat{\mu}$ is the same for both GMM stages, the estimate is reported once. The variables $R_t^e$ and $f_t$ are defined in the note to Table 1. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^\diamond$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-HAC standard errors are reported in parentheses for $\hat{\mu}$, $\hat{b}^\diamond$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns, given by $d_T \hat{b}^\diamond$, where $d_T = \frac{1}{T}\sum_{t=1}^{T} R_t^e (f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 2b: **GMM Estimates of Linear Factor Models**
**Lustig-Verdelhan Dataset, $\xi$-Normalization, Hybrid Weighting Matrix**

| | First Stage | | | | Second Stage | | | |
|---|---|---|---|---|---|---|---|---|
| | $\mu$ | $b^\diamond$ | $\lambda_f$ | $R^2$ | $b^\diamond$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | 0.016 | 45.1 | 0.95 | 0.10 | 89.4 | 1.88 | 0.00 | 15.1 |
| | (0.003) | (74.7) | (1.48) | | (30.0) | (0.81) | | (0.035) |
| **Durables growth** | 0.034 | 20.8 | 0.87 | 0.16 | 68.1 | 2.85 | -0.71 | 15.5 |
| | (0.007) | (29.2) | (1.15) | | (21.3) | (0.77) | | (0.030) |
| ***Rm-Rf*** | 0.070 | 1.8 | 5.97 | 0.02 | 2.4 | 7.69 | 0.02 | 20.2 |
| | (0.025) | (3.6) | (11.83) | | (2.4) | (7.70) | | (0.005) |
| **Consumption Factors** | | | | | | | | |
| Consumption growth | 0.016 | -8.7 | 0.26 | 0.16 | 1.6 | 1.27 | -0.61 | 16.1 |
| | (0.003) | (56.9) | (0.80) | | (40.5) | (0.73) | | (0.013) |
| Durables growth | 0.034 | 23.6 | 0.82 | | 64.9 | 2.74 | | |
| | (0.007) | (39.0) | (1.05) | | (27.6) | (0.78) | | |
| **Yogo Factors** | | | | | | | | |
| Consumption growth | 0.016 | -22.0 | 0.59 | 0.34 | 33.0 | 2.51 | -0.84 | 9.5 |
| | (0.003) | (63.6) | (1.18) | | (45.3) | (1.08) | | (0.089) |
| Durables growth | 0.034 | 45.5 | 1.10 | | 84.2 | 3.71 | | |
| | (0.007) | (51.0) | (1.78) | | (33.1) | (1.24) | | |
| *Rm* | 0.070 | 5.2 | 11.74 | | 5.8 | 12.85 | | |
| | (0.025) | (3.0) | (9.42) | | (2.8) | (8.61) | | |
| **Fama-French Factors** | | | | | | | | |
| *Rm-Rf* | 0.070 | 1.5 | 7.07 | 0.08 | -0.2 | 2.32 | 0.06 | 15.6 |
| | (0.025) | (4.3) | (11.36) | | (3.5) | (9.65) | | (0.008) |
| *SMB* | 0.024 | 1.7 | 4.08 | | 2.8 | 5.02 | | |
| | (0.020) | (4.6) | (7.06) | | (3.9) | (6.04) | | |
| *HML* | 0.057 | -2.8 | -5.91 | | -3.6 | -6.97 | | |
| | (0.020) | (5.1) | (8.78) | | (4.0) | (7.36) | | |

*Note*: Annual data, 1953–2002. The table reports first and second stage GMM estimates of $\mu$ and $b^\diamond$, from the SDF $m_t = 1 - (f_t - \mu)' b^\diamond$, obtained using the moment restrictions $E(R^e_t m_t) = 0$, $E(f_t - \mu) = 0$. The weighting matrix at the second stage of GMM uses the hybrid approach described in the main text. Since $\hat{\mu}$ is the same for both GMM stages, the estimate is reported once. The variables $R^e_t$ and $f_t$ are defined in the note to Table 1. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^\diamond$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-VARHAC standard errors are reported in parentheses for $\hat{\mu}$, $\hat{b}^\diamond$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R^e_t$ and the predicted mean returns, given by $d_T \hat{b}^\diamond$, where $d_T = \frac{1}{T}\sum_{t=1}^{T} R^e_t (f'_t - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 2c: **GMM Estimates of Linear Factor Models**
**Lustig-Verdelhan Dataset, $\xi$-Normalization, Cochrane's Weighting Matrix**

| | First Stage | | | | Second Stage | | | |
|---|---|---|---|---|---|---|---|---|
| | $\mu$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | 0.016 | 45.1 | 0.95 | 0.10 | 89.4 | 1.88 | 0.00 | 15.1 |
| | (0.003) | (74.7) | (1.48) | | (30.0) | (0.81) | | (0.035) |
| **Durables growth** | 0.034 | 20.8 | 0.87 | 0.16 | 68.1 | 2.85 | -0.71 | 15.5 |
| | (0.007) | (29.2) | (1.15) | | (21.3) | (0.77) | | (0.030) |
| ***Rm-Rf*** | 0.070 | 1.8 | 5.97 | 0.02 | 2.4 | 7.69 | 0.02 | 20.2 |
| | (0.025) | (3.6) | (11.83) | | (2.4) | (7.70) | | (0.005) |
| **Consumption Factors** | | | | | | | | |
| Consumption growth | 0.016 | -8.7 | 0.26 | 0.16 | 1.6 | 1.27 | -0.61 | 16.1 |
| | (0.003) | (56.9) | (0.80) | | (40.5) | (0.73) | | (0.013) |
| Durables growth | 0.034 | 23.6 | 0.82 | | 64.9 | 2.74 | | |
| | (0.007) | (39.0) | (1.05) | | (27.6) | (0.78) | | |
| **Yogo Factors** | | | | | | | | |
| Consumption growth | 0.016 | -22.0 | 0.59 | 0.34 | 33.0 | 2.51 | -0.84 | 9.5 |
| | (0.003) | (63.6) | (1.18) | | (45.3) | (1.08) | | (0.089) |
| Durables growth | 0.034 | 45.5 | 1.10 | | 84.2 | 3.71 | | |
| | (0.007) | (51.0) | (1.78) | | (33.1) | (1.24) | | |
| *Rm* | 0.070 | 5.2 | 11.74 | | 5.8 | 12.85 | | |
| | (0.025) | (3.0) | (9.42) | | (2.8) | (8.61) | | |
| **Fama-French Factors** | | | | | | | | |
| *Rm-Rf* | 0.070 | 1.5 | 7.07 | 0.08 | -0.2 | 2.32 | 0.06 | 15.6 |
| | (0.025) | (4.3) | (11.36) | | (3.5) | (9.65) | | (0.008) |
| *SMB* | 0.024 | 1.7 | 4.08 | | 2.8 | 5.02 | | |
| | (0.020) | (4.6) | (7.06) | | (3.9) | (6.04) | | |
| *HML* | 0.057 | -2.8 | -5.91 | | -3.6 | -6.97 | | |
| | (0.020) | (5.1) | (8.78) | | (4.0) | (7.36) | | |

*Note*: Annual data, 1953–2002. The table reports first and second stage GMM estimates of $\mu$ and $b^{\diamond}$, from the SDF $m_t = 1 - (f_t - \mu)'b^{\diamond}$, obtained using the moment restrictions $E(R_t^e m_t) = 0$, $E(f_t - \mu) = 0$. The weighting matrix at the second stage of GMM uses Cochrane's approach, described in the main text. Since $\hat{\mu}$ is the same for both GMM stages, the estimate is reported once. The variables $R_t^e$ and $f_t$ are defined in the note to Table 1. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^{\diamond}$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-VARHAC standard errors are reported in parentheses for $\hat{\mu}$, $\hat{b}^{\diamond}$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns, given by $d_T \hat{b}^{\diamond}$, where $d_T = \frac{1}{T}\sum_{t=1}^{T} R_t^e (f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 2d: **GMM Estimates of Linear Factor Models**
**Lustig-Verdelhan Dataset, $\xi$-Normalization, Model has a Common Pricing Error Parameter**

| | First Stage | | | | Second Stage | | | |
|---|---|---|---|---|---|---|---|---|
| | $\mu$ | $b^\diamond$ | $\lambda_f$ | $R^2$ | $b^\diamond$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | | | | | | | | |
| Pricing Error ($\gamma$) | | | -0.69 | 0.18 | | -3.28 | -1.79 | 5.1 |
| | | | (1.81) | | | (1.20) | | (0.536) |
| Consumption growth | 0.016 | 92.0 | 1.94 | | 137.3 | 2.89 | | |
| | (0.003) | (60.1) | (1.24) | | (46.3) | (1.04) | | |
| **Durables growth** | | | | | | | | |
| Pricing Error ($\gamma$) | | | -3.06 | 0.74 | | -1.89 | 0.54 | 1.4 |
| | | | (2.73) | | | (1.89) | | (0.965) |
| Durables growth | 0.034 | 111.4 | 4.65 | | 59.8 | 2.50 | | |
| | (0.007) | (88.8) | (3.17) | | (72.5) | (2.75) | | |
| **Market premium** | | | | | | | | |
| Pricing Error ($\gamma$) | | | 0.25 | 0.04 | | -1.36 | -0.95 | 19.6 |
| | | | (1.07) | | | (0.92) | | (0.003) |
| *Rm-Rf* | 0.070 | 2.4 | 7.92 | | 1.5 | 4.83 | | |
| | (0.025) | (2.9) | (9.50) | | (2.5) | (8.05) | | |
| **Consumption Factors** | | | | | | | | |
| Pricing Error ($\gamma$) | | | -3.06 | 0.74 | | -2.83 | 0.45 | 1.0 |
| | | | (2.74) | | | (2.40) | | (0.965) |
| Consumption growth | 0.016 | -9.5 | 1.97 | | 42.5 | 2.16 | | |
| | (0.003) | (88.0) | (2.06) | | (65.5) | (2.02) | | |
| Durables growth | 0.034 | 114.4 | 4.60 | | 66.5 | 3.59 | | |
| | (0.007) | (93.8) | (3.26) | | (73.0) | (3.13) | | |
| **Yogo Factors** | | | | | | | | |
| Pricing Error ($\gamma$) | | | -2.94 | 0.87 | | -1.91 | 0.41 | 1.2 |
| | | | (2.92) | | | (2.29) | | (0.886) |
| Consumption growth | 0.016 | -21.0 | 2.19 | | 6.9 | 1.22 | | |
| | (0.003) | (88.6) | (2.09) | | (76.4) | (2.00) | | |
| Durables growth | 0.034 | 129.9 | 4.70 | | 52.6 | 2.18 | | |
| | (0.007) | (109.5) | (3.63) | | (77.5) | (3.12) | | |
| *Rm* | 0.070 | 4.5 | 3.33 | | 2.0 | 2.55 | | |
| | (0.025) | (5.1) | (13.30) | | (4.0) | (11.18) | | |
| **Fama-French Factors** | | | | | | | | |
| Pricing Error ($\gamma$) | | | -0.28 | 0.09 | | -1.05 | -0.30 | 15.7 |
| | | | (1.00) | | | (0.95) | | (0.004) |
| *Rm-Rf* | 0.070 | 1.1 | 5.72 | | -0.9 | -0.59 | | |
| | (0.025) | (4.3) | (11.24) | | (3.4) | (9.32) | | |
| *SMB* | 0.024 | 1.5 | 3.50 | | 2.1 | 3.40 | | |
| | (0.020) | (4.6) | (6.58) | | (3.9) | (6.00) | | |
| *HML* | 0.057 | -3.6 | -7.26 | | -2.8 | -5.14 | | |
| | (0.020) | (4.5) | (7.85) | | (4.0) | (7.14) | | |

The note to Table 2d is on the following page.

*Note to Table 2d*: Annual data, 1953–2002. The table reports first and second stage GMM estimates of $\mu$ and $b^{\diamond}$, from the SDF $m_t = 1 - (f_t - \mu)' b^{\diamond}$, obtained using the moment restrictions $E(R_t^e m_t) = \gamma$, $E(f_t - \mu) = 0$. Since $\hat{\mu}$ is the same for both GMM stages, the estimate is reported once. The variables $R_t^e$ and $f_t$ are defined in the note to Table 1. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^{\diamond}$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-VAR$a$ HAC standard errors are reported in parentheses for $\hat{\mu}$, $\hat{b}^{\diamond}$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns, given by $d_T \hat{b}^{\diamond}$, where $d_T = \frac{1}{T} \sum_{t=1}^{T} R_t^e (f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 2e: **GMM Estimates of Linear Factor Models**
**Lustig-Verdelhan Dataset, $\xi$-Normalization, Using all Information to Estimate $\mu$**

| | First Stage | | | | Second Stage | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\mu$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $\mu$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | 0.011 | 72.8 | 1.66 | 0.49 | 0.011 | 103.8 | 2.36 | 0.38 | 14.7 |
| | (0.003) | (60.4) | (1.61) | | (0.002) | (27.3) | (1.10) | | (0.039) |
| **Durables growth** | 0.029 | 26.0 | 1.34 | 0.31 | 0.024 | 49.8 | 2.56 | 0.18 | 10.0 |
| | (0.007) | (25.1) | (1.35) | | (0.006) | (18.7) | (1.17) | | (0.190) |
| ***Rm-Rf*** | 0.010 | 5.4 | 20.03 | 0.40 | 0.136 | -4.1 | -15.13 | 0.23 | 19.3 |
| | (0.036) | (2.9) | (12.20) | | (0.019) | (2.6) | (8.54) | | (0.007) |
| **Consumption Factors** | | | | | | | | | |
| Consumption growth | 0.011 | 106.4 | 1.95 | 0.55 | 0.014 | 168.4 | 2.21 | 0.18 | 12.1 |
| | (0.003) | (79.2) | (1.17) | | (0.003) | (52.3) | (1.33) | | (0.059) |
| Durables growth | 0.030 | -18.1 | 1.08 | | 0.040 | -77.6 | -0.51 | | |
| | (0.007) | (68.0) | (2.24) | | (0.006) | (47.4) | (1.93) | | |
| **Yogo Factors** | | | | | | | | | |
| Consumption growth | 0.013 | 22.3 | 1.27 | 0.63 | 0.013 | 23.0 | 1.97 | -0.19 | 5.9 |
| | (0.003) | (67.1) | (1.35) | | (0.003) | (47.2) | (1.24) | | (0.312) |
| Durables growth | 0.031 | 28.9 | 1.35 | | 0.027 | 66.1 | 3.28 | | |
| | (0.007) | (57.6) | (2.21) | | (0.005) | (36.8) | (1.80) | | |
| *Rm* | 0.036 | 5.6 | 15.82 | | 0.080 | 3.5 | 6.35 | | |
| | (0.031) | (2.7) | (15.22) | | (0.024) | (2.6) | (11.56) | | |
| **Fama-French Factors** | | | | | | | | | |
| *Rm-Rf* | 0.059 | -0.8 | 4.64 | 0.61 | 0.082 | -4.6 | -5.71 | 0.40 | 11.9 |
| | (0.025) | (4.8) | (15.72) | | (0.023) | (3.4) | (12.53) | | (0.037) |
| *SMB* | -0.008 | 8.4 | 15.76 | | -0.006 | 11.1 | 18.96 | | |
| | (0.028) | (4.2) | (9.96) | | (0.022) | (3.0) | (9.73) | | |
| *HML* | 0.089 | -7.6 | -14.68 | | 0.066 | -9.4 | -17.01 | | |
| | (0.022) | (3.9) | (8.50) | | (0.018) | (3.3) | (9.23) | | |

*Note*: Annual data, 1953–2002. The table reports first and second stage GMM estimates of $\mu$ and $b^{\diamond}$, from the SDF $m_t = 1 - (f_t - \mu)' b^{\diamond}$, obtained using the moment restrictions $E(R^e_t m_t) = 0$, $E(f_t - \mu) = 0$. The weighting matrices at the two stages of GMM are based on Yogo (2006) and are described in the main text. The variables $R^e_t$ and $f_t$ are defined in the note to Table 1. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^{\diamond}$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-VARHAC standard errors are reported in parentheses for $\hat{\mu}$, $\hat{b}^{\diamond}$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R^e_t$ and the predicted mean returns, given by $d_T \hat{b}^{\diamond}$, where $d_T = \frac{1}{T}\sum_{t=1}^{T} R^e_t (f'_t - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 1f/2f: **GMM Estimates of Linear Factor Models**
**Lustig-Verdelhan Dataset,** $a$ and $\xi$**-Normalizations, Weighting Matrices Iterated to Convergence**

| | a-Normalization | | | | $\xi$-Normalization | | | |
|---|---|---|---|---|---|---|---|---|
| | $b^*$ | $\lambda_f^*$ | $R^2$ | $J$ | $b^\diamond$ | $\lambda_f^\diamond$ | $R^2$ | $J$ |
| **Consumption growth** | 48.8 | 4.29 | 0.87 | 5.1 | 91.5 | 1.93 | -0.01 | 11.9 |
| | (4.3) | (1.64) | | (0.647) | (42.7) | (0.80) | | (0.105) |
| **Durables growth** | 24.4 | 5.70 | 0.81 | 4.3 | 36.1 | 1.51 | 0.07 | 11.3 |
| | (1.9) | (4.50) | | (0.747) | (26.5) | (0.95) | | (0.127) |
| ***Rm-Rf*** | 6.7 | 41.49 | 0.44 | 8.7 | 0.9 | 3.10 | 0.02 | 21.5 |
| | (1.9) | (20.5) | | (0.275) | (2.3) | (7.48) | | (0.003) |
| **Consumption Factors** | | | | | | | | |
| Consumption growth | 0.6 | 2.61 | 0.81 | 4.3 | -20.5 | 0.37 | 0.08 | 10.5 |
| | (10.5) | (2.09) | | (0.633) | (39.7) | (0.71) | | (0.104) |
| Durables growth | 24.1 | 5.64 | | | 42.1 | 1.37 | | |
| | (5.4) | (4.37) | | | (30.5) | (0.98) | | |
| **Yogo Factors** | | | | | | | | |
| Consumption growth | -4.0 | 2.49 | 0.85 | 4.9 | -44.4 | 0.17 | 0.17 | 7.8 |
| | (12.6) | (2.31) | | (0.425) | (49.8) | (0.79) | | (0.170) |
| Durables growth | 25.2 | 5.50 | | | 54.7 | 1.32 | | |
| | (5.9) | (4.73) | | | (35.3) | (1.09) | | |
| *Rm* | 0.7 | 0.2 | | | 1.7 | -0.51 | | |
| | (0.7) | (12.84) | | | (2.0) | (5.70) | | |
| **Fama-French Factors** | | | | | | | | |
| *Rm-Rf* | 7.3 | 80.19 | 0.28 | 3.9 | -1.7 | -3.20 | 0.00 | 15.3 |
| | (3.2) | (85.7) | | (0.557) | (3.3) | (8.76) | | (0.009) |
| *SMB* | -3.9 | -11.84 | | | 2.4 | 3.32 | | |
| | (4.5) | (33.66) | | | (3.5) | (5.18) | | |
| *HML* | 6.1 | 37.69 | | | -3.0 | -5.19 | | |
| | (2.8) | (40.05) | | | (3.5) | (6.27) | | |

*Note*: Annual data, 1953–2002. The reports GMM stimates of $b^*$, from the SDF $m_t = 1 - f_t' b^*$, obtained using the moment restriction $E(R_t^e m_t) = 0$ and GMM estimates of $\mu$ and $b^\diamond$, from the SDF $m_t = 1 - (f_t - \mu)' b^\diamond$, obtained using the moment restrictions $E(R_t^e m_t) = 0$, $E(f_t - \mu) = 0$. Both sets of estimates are based on iterating the weighting matrix to convergence. The variables $R_t^e$ and $f_t$ are defined in the note to Table 1. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^\diamond$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-VARHAC standard errors are reported in parentheses for $b^*$, $\hat{\mu}$, $\hat{b}^\diamond$, $\hat{\lambda}_f^*$ and $\hat{\lambda}_f^\diamond$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns. For the $a$-normalization the predicted mean returns are $D_T \hat{b}^*$, where $D_T = \frac{1}{T}\sum_{t=1}^T R_t^e f_t'$. For the $\xi$-normalization they are $d_T \hat{b}^\diamond$, where $d_T = \frac{1}{T}\sum_{t=1}^T R_t^e (f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 3a: **GMM Estimates of Linear Factor Models**
**Fama-French 25 Dataset, $a$-Normalization, HAC Standard Errors**

| | First Stage | | | Second Stage | | | |
|---|---|---|---|---|---|---|---|
| | $b^*$ | $\lambda_f^*$ | $R^2$ | $b^*$ | $\lambda_f^*$ | $R^2$ | $J$ |
| **Consumption growth** | 126.3 | 0.94 | 0.81 | 142.8 | 1.38 | 0.51 | 30.2 |
| | (23.6) | (0.50) | | (10.1) | (0.33) | | (0.179) |
| **Durables growth** | 111.1 | -2.48 | 0.89 | 93.3 | 11.03 | 0.43 | 15.5 |
| | (18.3) | (2.5) | | (7.2) | (25.1) | | (0.906) |
| ***Rm-Rf*** | 3.3 | 2.31 | -0.55 | 4.7 | 3.40 | -3.55 | 65.5 |
| | (0.9) | (0.57) | | (0.8) | (0.59) | | (0.000) |
| **Consumption Factors** | | | | | | | |
| Consumption growth | 54.0 | 2.83 | 0.98 | 49.3 | 1.93 | 0.96 | 17.2 |
| | (33.8) | (3.69) | | (13.8) | (1.15) | | (0.800) |
| Durables growth | 64.4 | 4.24 | | 64.8 | 3.15 | | |
| | (20.4) | (6.64) | | (8.3) | (1.92) | | |
| **Yogo Factors** | | | | | | | |
| Consumption growth | 41.0 | 2.10 | 0.98 | 38.5 | 1.47 | 0.97 | 17.3 |
| | (32.5) | (3.16) | | (16.6) | (0.99) | | (0.748) |
| Durables growth | 69.9 | 4.21 | | 68.7 | 3.02 | | |
| | (18.1) | (6.29) | | (8.8) | (1.78) | | |
| *Rm* | 0.19 | 2.33 | | 0.36 | 2.96 | | |
| | (0.53) | (1.45) | | (0.44) | (1.40) | | |
| **Fama-French Factors** | | | | | | | |
| *Rm-Rf* | 3.9 | 1.92 | 0.75 | 5.0 | 2.53 | -0.20 | 47.8 |
| | (0.9) | (0.54) | | (0.8) | (0.56) | | (0.001) |
| *SMB* | -0.1 | 0.52 | | -0.5 | 0.58 | | |
| | (1.2) | (0.36) | | (1.1) | (0.37) | | |
| *HML* | 5.9 | 1.35 | | 7.2 | 1.69 | | |
| | (1.1) | (0.37) | | (0.9) | (0.40) | | |

*Note*: Quarterly data, 1949–2005. The table reports first and second stage GMM estimates of $b^*$, from the SDF $m_t = 1 - f_t' b^*$, obtained using the moment restriction $E(R_t^e m_t) = 0$, where $R_t^e$ is a $25 \times 1$ vector of excess returns of the Fama-French 25 portfolios of US stocks sorted on size and the book-to-market value ratio, and $f_t$ is a scalar or vector of risk factors. The factors are real per capita consumption (nondurables & services) growth, real per capita durable consumption growth, and the following variables from the Fama-French dataset: the real value weighted US stock market excess return over the risk free rate (*Rm-Rf*), the gross return to the same portfolio (*Rm*), and the *SMB* and *HML* portfolio excess returns (see Appendix B). Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^* / (1 - \bar{f}' \hat{b}^*)$ are also reported (in percent), where $\bar{f}$ and $\mathrm{S}_f$ are the sample mean and covariance matrix of $f_t$. GMM-HAC standard errors are reported in parentheses for $\hat{b}^*$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns, given by $D_T \hat{b}^*$, where $D_T = \frac{1}{T} \sum_{t=1}^{T} R_t^e f_t'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 4a: **GMM Estimates of Linear Factor Models**
**Fama-French 25 Dataset, $\xi$-Normalization, HAC Standard Errors**

| | First Stage | | | | Second Stage | | | |
|---|---|---|---|---|---|---|---|---|
| | $\mu$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | 0.0051 | 335.4 | 0.89 | -0.44 | 173.1 | 0.46 | -4.30 | 39.1 |
| | (0.0003) | (168.0) | (0.47) | | (69.8) | (0.17) | | (0.026) |
| **Durables growth** | 0.0104 | -554.0 | -1.92 | -2.46 | -64.4 | -0.22 | -13.77 | 10.0 |
| | (0.0004) | (466.2) | (1.7) | | (123.6) | (0.4) | | (0.994) |
| ***Rm-Rf*** | 0.0197 | 3.5 | 2.30 | -0.77 | 3.1 | 2.02 | -1.00 | 69.8 |
| | (0.0053) | (1.0) | (0.57) | | (1.0) | (0.54) | | (0.000) |
| **Consumption Factors** | | | | | | | | |
| Consumption growth | 0.0051 | 370.3 | 0.99 | -0.42 | 185.5 | 0.49 | -4.19 | 33.7 |
| | (0.0003) | (161.8) | (0.45) | | (80.0) | (0.20) | | (0.069) |
| Durables growth | 0.0104 | 64.9 | 0.31 | | 18.8 | 0.11 | | |
| | (0.0004) | (163.5) | (0.55) | | (74.5) | (0.25) | | |
| **Yogo Factors** | | | | | | | | |
| Consumption growth | 0.0051 | 271.6 | 0.76 | -0.33 | 66.0 | 0.22 | -2.28 | 31.2 |
| | (0.0003) | (168.8) | (0.42) | | (82.0) | (0.21) | | (0.092) |
| Durables growth | 0.0104 | 136.4 | 0.53 | | 88.4 | 0.31 | | |
| | (0.0004) | (131.2) | (0.44) | | (62.4) | (0.22) | | |
| *Rm* | 0.0224 | 1.5 | 2.18 | | 4.5 | 3.05 | | |
| | (0.0053) | (2.4) | (0.63) | | (1.7) | (0.70) | | |
| **Fama-French Factors** | | | | | | | | |
| *Rm-Rf* | 0.0197 | 4.5 | 1.92 | 0.66 | 4.4 | 1.77 | 0.55 | 52.7 |
| | (0.0053) | (1.2) | (0.54) | | (1.1) | (0.54) | | (0.000) |
| *SMB* | 0.0063 | -0.1 | 0.53 | | -0.5 | 0.40 | | |
| | (0.0035) | (1.4) | (0.36) | | (1.3) | (0.35) | | |
| *HML* | 0.0119 | 6.8 | 1.32 | | 7.0 | 1.40 | | |
| | (0.0036) | (1.4) | (0.37) | | (1.3) | (0.36) | | |

*Note*: Quarterly data, 1949–2005. The table reports first and second stage GMM estimates of $\mu$ and $b^{\diamond}$, from the SDF $m_t = 1-(f_t-\mu)'b^{\diamond}$, obtained using the moment restrictions $E(R^e_t m_t)=0$, $E(f_t-\mu)=0$. Since $\hat{\mu}$ is the same for both GMM stages, the estimate is reported once. The variables $R^e_t$ and $f_t$ are defined in the note to Table 3. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f\hat{b}^{\diamond}$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-HAC standard errors are reported in parentheses for $\hat{\mu}$, $\hat{b}^{\diamond}$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R^e_t$ and the predicted mean returns, given by $d_T\hat{b}^{\diamond}$, where $d_T = \frac{1}{T}\sum_{t=1}^{T} R^e_t(f'_t-\hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 4b: **GMM Estimates of Linear Factor Models**
**Fama-French 25 Dataset, $\xi$-Normalization, Hybrid Weighting Matrix**

| | First Stage | | | | Second Stage | | | |
|---|---|---|---|---|---|---|---|---|
| | $\mu$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | 0.0051 | 335.4 | 0.89 | -0.44 | 282.4 | 0.75 | -0.85 | 27.4 |
| | (0.0005) | (164.5) | (0.46) | | (77.0) | (0.19) | | (0.284) |
| **Durables growth** | 0.0104 | -554.0 | -1.92 | -2.46 | -310.3 | -1.08 | -5.26 | 7.3 |
| | (0.0012) | (534.1) | (1.9) | | (332.2) | (1.2) | | (1.000) |
| ***Rm-Rf*** | 0.0197 | 3.5 | 2.30 | -0.77 | 4.3 | 2.80 | -1.54 | 68.3 |
| | (0.0054) | (1.1) | (0.57) | | (1.0) | (0.53) | | (0.000) |
| **Consumption Factors** | | | | | | | | |
| Consumption growth | 0.0051 | 370.3 | 0.99 | -0.42 | 314.2 | 0.84 | -0.75 | 22.9 |
| | (0.0005) | (166.6) | (0.46) | | (93.2) | (0.23) | | (0.468) |
| Durables growth | 0.0104 | 64.9 | 0.31 | | 48.5 | 0.24 | | |
| | (0.0012) | (165.2) | (0.56) | | (74.6) | (0.25) | | |
| **Yogo Factors** | | | | | | | | |
| Consumption growth | 0.0051 | 271.6 | 0.76 | -0.33 | 176.8 | 0.52 | -2.91 | 23.8 |
| | (0.0005) | (177.6) | (0.45) | | (97.6) | (0.24) | | (0.356) |
| Durables growth | 0.0104 | 136.4 | 0.53 | | 119.8 | 0.44 | | |
| | (0.0012) | (136.2) | (0.48) | | (66.8) | (0.27) | | |
| *Rm* | 0.0224 | 1.5 | 2.18 | | 3.8 | 3.15 | | |
| | (0.0054) | (2.4) | (0.73) | | (1.8) | (0.70) | | |
| **Fama-French Factors** | | | | | | | | |
| *Rm-Rf* | 0.0197 | 4.5 | 1.92 | 0.66 | 5.3 | 2.16 | 0.46 | 52.1 |
| | (0.0054) | (1.2) | (0.63) | | (1.1) | (0.65) | | (0.000) |
| *SMB* | 0.0063 | -0.1 | 0.53 | | -0.8 | 0.44 | | |
| | (0.0036) | (1.4) | (0.36) | | (1.3) | (0.35) | | |
| *HML* | 0.0119 | 6.8 | 1.32 | | 8.0 | 1.56 | | |
| | (0.0036) | (1.4) | (0.49) | | (1.3) | (0.57) | | |

*Note*: Quarterly data, 1949–2005. The table reports first and second stage GMM estimates of $\mu$ and $b^{\diamond}$, from the SDF $m_t = 1 - (f_t - \mu)'b^{\diamond}$, obtained using the moment restrictions $E(R_t^e m_t) = 0$, $E(f_t - \mu) = 0$. The weighting matrix at the second stage of GMM uses the hybrid approach described in the main text. Since $\hat{\mu}$ is the same for both GMM stages, the estimate is reported once. The variables $R_t^e$ and $f_t$ are defined in the note to Table 3. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^{\diamond}$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-VARHAC standard errors are reported in parentheses for $\hat{\mu}$, $\hat{b}^{\diamond}$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns, given by $d_T \hat{b}^{\diamond}$, where $d_T = \frac{1}{T}\sum_{t=1}^{T} R_t^e (f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 4c: **GMM Estimates of Linear Factor Models**
**Fama-French 25 Dataset, $\xi$-Normalization, Cochrane's Weighting Matrix**

| | First Stage | | | | Second Stage | | | |
|---|---|---|---|---|---|---|---|---|
| | $\mu$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | 0.0051 | 335.4 | 0.89 | -0.44 | 150.8 | 0.40 | -5.43 | 26.6 |
| | (0.0005) | (164.5) | (0.46) | | (84.8) | (0.20) | | (0.324) |
| **Durables growth** | 0.0104 | -554.0 | -1.92 | -2.46 | -67.9 | -0.24 | -13.60 | 6.8 |
| | (0.0012) | (534.1) | (1.9) | | (400.4) | (1.4) | | (1.000) |
| ***Rm-Rf*** | 0.0197 | 3.5 | 2.30 | -0.77 | 4.2 | 2.73 | -1.35 | 68.4 |
| | (0.0054) | (1.1) | (0.57) | | (1.0) | (0.53) | | (0.000) |
| **Consumption Factors** | | | | | | | | |
| Consumption growth | 0.0051 | 370.3 | 0.99 | -0.42 | 156.0 | 0.41 | -5.44 | 22.5 |
| | (0.0005) | (166.6) | (0.46) | | (102.8) | (0.25) | | (0.489) |
| Durables growth | 0.0104 | 64.9 | 0.31 | | 10.0 | 0.07 | | |
| | (0.0012) | (165.2) | (0.56) | | (75.8) | (0.26) | | |
| **Yogo Factors** | | | | | | | | |
| Consumption growth | 0.0051 | 271.6 | 0.76 | -0.33 | 85.9 | 0.26 | -0.55 | 24.0 |
| | (0.0005) | (177.6) | (0.45) | | (101.3) | (0.26) | | (0.348) |
| Durables growth | 0.0104 | 136.4 | 0.53 | | 72.6 | 0.26 | | |
| | (0.0012) | (136.2) | (0.48) | | (68.2) | (0.26) | | |
| *Rm* | 0.0224 | 1.5 | 2.18 | | 3.2 | 2.37 | | |
| | (0.0054) | (2.4) | (0.73) | | (1.8) | (0.78) | | |
| **Fama-French Factors** | | | | | | | | |
| *Rm-Rf* | 0.0197 | 4.5 | 1.92 | 0.66 | 5.2 | 2.11 | 0.51 | 50.9 |
| | (0.0054) | (1.2) | (0.63) | | (1.1) | (0.65) | | (0.000) |
| *SMB* | 0.0063 | -0.1 | 0.53 | | -0.6 | 0.46 | | |
| | (0.0036) | (1.4) | (0.36) | | (1.3) | (0.35) | | |
| *HML* | 0.0119 | 6.8 | 1.32 | | 7.9 | 1.54 | | |
| | (0.0036) | (1.4) | (0.49) | | (1.3) | (0.56) | | |

*Note*: Quarterly data, 1949–2005. The table reports first and second stage GMM estimates of $\mu$ and $b^{\diamond}$, from the SDF $m_t = 1 - (f_t - \mu)' b^{\diamond}$, obtained using the moment restrictions $E(R_t^e m_t) = 0$, $E(f_t - \mu) = 0$. The weighting matrix at the second stage of GMM uses Cochrane's approach, described in the main text. Since $\hat{\mu}$ is the same for both GMM stages, the estimate is reported once. The variables $R_t^e$ and $f_t$ are defined in the note to Table 3. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^{\diamond}$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-VARHAC standard errors are reported in parentheses for $\hat{\mu}$, $\hat{b}^{\diamond}$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns, given by $d_T \hat{b}^{\diamond}$, where $d_T = \frac{1}{T}\sum_{t=1}^{T} R_t^e (f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 4d: **GMM Estimates of Linear Factor Models**
**Fama-French 25 Dataset, $\xi$-Normalization, Model has a Common Pricing Error Parameter**

| | First Stage | | | | Second Stage | | | |
|---|---|---|---|---|---|---|---|---|
| | $\mu$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $b^{\diamond}$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | | | | | | | | |
| Pricing Error (?) | | | 1.85 | 0.10 | | 2.50 | -0.05 | 52.6 |
| | | | (0.63) | | | (0.45) | | (0.000) |
| Consumption growth | 0.0051 | 104.1 | 0.28 | | 47.9 | 0.13 | | |
| | (0.0005) | (91.3) | (0.24) | | (41.7) | (0.11) | | |
| **Durables growth** | | | | | | | | |
| Pricing Error (?) | | | 2.68 | 0.00 | | 2.57 | -0.07 | 55.0 |
| | | | (0.49) | | | (0.41) | | (0.000) |
| Durables growth | 0.0104 | 13.2 | 0.05 | | 25.7 | 0.09 | | |
| | (0.0012) | (92.1) | (0.32) | | (32.1) | (0.12) | | |
| **Market premium** | | | | | | | | |
| Pricing Error (?) | | | 3.52 | 0.05 | | 3.28 | -0.58 | 55.7 |
| | | | (0.90) | | | (0.62) | | (0.000) |
| *Rm-Rf* | 0.0197 | -1.2 | -0.81 | | -1.6 | -1.05 | | |
| | (0.0054) | (1.6) | (1.04) | | (1.2) | (0.82) | | |
| **Consumption Factors** | | | | | | | | |
| Pricing Error (?) | | | 1.93 | 0.16 | | 1.91 | -1.35 | 24.6 |
| | | | (0.85) | | | (0.61) | | (0.316) |
| Consumption growth | 0.0051 | 153.9 | 0.43 | | 17.7 | 0.06 | | |
| | (0.0005) | (106.8) | (0.28) | | (60.0) | (0.16) | | |
| Durables growth | 0.0104 | 111.6 | 0.42 | | 40.7 | 0.15 | | |
| | (0.0012) | (110.5) | (0.39) | | (43.9) | (0.16) | | |
| **Yogo Factors** | | | | | | | | |
| Pricing Error (?) | | | 3.78 | 0.45 | | 3.08 | -3.36 | 28.8 |
| | | | (1.14) | | | (0.88) | | (0.119) |
| Consumption growth | 0.0051 | 253.3 | 0.63 | | 49.5 | 0.12 | | |
| | (0.0005) | (153.1) | (0.41) | | (70.0) | (0.17) | | |
| Durables growth | 0.0104 | -65.3 | -0.15 | | 7.9 | 0.05 | | |
| | (0.0012) | (96.7) | (0.27) | | (57.1) | (0.20) | | |
| *Rm* | 0.0224 | -4.6 | -1.24 | | -2.8 | -1.53 | | |
| | (0.0054) | (2.4) | (1.30) | | (2.1) | (1.17) | | |
| **Fama-French Factors** | | | | | | | | |
| Pricing Error (?) | | | 3.83 | 0.77 | | 3.36 | 0.71 | 44.5 |
| | | | (1.05) | | | (0.77) | | (0.002) |
| *Rm-Rf* | 0.0197 | -3.2 | -1.82 | | -2.3 | -1.36 | | |
| | (0.0054) | (2.2) | (1.16) | | (1.7) | (0.95) | | |
| *SMB* | 0.0063 | 4.2 | 0.49 | | 4.3 | 0.65 | | |
| | (0.0036) | (1.7) | (0.37) | | (1.6) | (0.37) | | |
| *HML* | 0.0119 | 3.2 | 1.25 | | 4.1 | 1.38 | | |
| | (0.0036) | (1.7) | (0.43) | | (1.5) | (0.47) | | |

The note to Table 4d is on the following page.

*Note to Table 4d*: Quarterly data, 1949–2005. The table reports first and second stage GMM estimates of $\mu$ and $b^{\diamond}$, from the SDF $m_t = 1 - (f_t - \mu)' b^{\diamond}$, obtained using the moment restrictions $E(R_t^e m_t) = \gamma$, $E(f_t - \mu) = 0$. Since $\hat{\mu}$ is the same for both GMM stages, the estimate is reported once. The variables $R_t^e$ and $f_t$ are defined in the note to Table 3. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^{\diamond}$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-VARHAC standard errors are reported in parentheses for $\hat{\mu}$, $\hat{b}^{\diamond}$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns, given by $d_T \hat{b}^{\diamond}$, where $d_T = \frac{1}{T}\sum_{t=1}^{T} R_t^e (f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 4e: **GMM Estimates of Linear Factor Models**
**Fama-French 25 Dataset, $\xi$-Normalization, Using all Information to Estimate $\mu$**

| | First Stage | | | | Second Stage | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\mu$ | $b^\diamond$ | $\lambda_f$ | $R^2$ | $\mu$ | $b^\diamond$ | $\lambda_f$ | $R^2$ | $J$ |
| **Consumption growth** | 0.0027 | 189.2 | 0.63 | 0.57 | 0.0025 | 240.0 | 0.80 | -1.10 | 67.0 |
| | (0.0008) | (60.7) | (0.21) | | (0.0005) | (29.5) | (0.15) | | (0.000) |
| **Durables growth** | 0.0136 | -207.8 | -1.22 | 0.60 | 0.0153 | -155.5 | -0.91 | 0.78 | 26.0 |
| | (0.0017) | (92.2) | (0.7) | | (0.0013) | (33.4) | (0.3) | | (0.354) |
| ***Rm-Rf*** | -0.0133 | 3.2 | 2.11 | -0.42 | 0.0099 | 3.4 | 2.21 | -0.66 | 81.5 |
| | (0.0089) | (0.9) | (0.71) | | (0.0067) | (0.9) | (0.70) | | (0.000) |
| **Consumption Factors** | | | | | | | | | |
| Consumption growth | 0.0034 | 204.7 | 0.56 | 0.74 | 0.0054 | 65.4 | 0.19 | -1.20 | 26.7 |
| | (0.0007) | (90.3) | (0.26) | | (0.0005) | (58.5) | (0.18) | | (0.271) |
| Durables growth | 0.0085 | 174.1 | 0.91 | | 0.0064 | 196.8 | 1.01 | | |
| | (0.0012) | (109.0) | (0.77) | | (0.0009) | (37.4) | (0.52) | | |
| **Yogo Factors** | | | | | | | | | |
| Consumption growth | 0.0037 | 170.1 | 0.48 | 0.74 | 0.0054 | 11.9 | 0.07 | 0.70 | 21.8 |
| | (0.0007) | (109.9) | (0.31) | | (0.0005) | (64.2) | (0.19) | | (0.471) |
| Durables growth | 0.0083 | 202.6 | 0.99 | | 0.0067 | 202.0 | 0.97 | | |
| | (0.0011) | (98.3) | (0.60) | | (0.0009) | (39.6) | (0.36) | | |
| *Rm* | 0.0203 | 0.5 | 0.70 | | 0.0226 | 2.2 | 0.82 | | |
| | (0.0053) | (1.8) | (1.62) | | (0.0050) | (1.4) | (1.27) | | |
| **Fama-French Factors** | | | | | | | | | |
| *Rm-Rf* | 0.0135 | 4.3 | 1.82 | 0.69 | 0.0177 | 4.5 | 1.83 | 0.64 | 58.0 |
| | (0.0054) | (1.1) | (0.89) | | (0.0052) | (1.0) | (0.88) | | (0.000) |
| *SMB* | 0.0046 | -0.1 | 0.50 | | 0.0054 | -0.5 | 0.42 | | |
| | (0.0036) | (1.4) | (0.44) | | (0.0035) | (1.3) | (0.41) | | |
| *HML* | 0.0076 | 6.5 | 1.26 | | 0.0098 | 6.7 | 1.32 | | |
| | (0.0042) | (1.3) | (0.72) | | (0.0041) | (1.2) | (0.70) | | |

*Note*: Quarterly data, 1949–2005. The table reports first and second stage GMM estimates of $\mu$ and $b^\diamond$, from the SDF $m_t = 1 - (f_t - \mu)'b^\diamond$, obtained using the moment restrictions $E(R_t^e m_t) = 0$, $E(f_t - \mu) = 0$. The weighting matrices at the two stages of GMM are based on Yogo (2006) and are described in the main text. The variables $R_t^e$ and $f_t$ are defined in the note to Table 3. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^\diamond$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-VARHAC standard errors are reported in parentheses for $\hat{\mu}$, $\hat{b}^\diamond$ and $\hat{\lambda}_f$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns, given by $d_T\hat{b}^\diamond$, where $d_T = \frac{1}{T}\sum_{t=1}^{T} R_t^e(f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.

TABLE 3f/4f: **GMM Estimates of Linear Factor Models**
**Fama-French 25 Dataset,** $a$ and $\xi$**-Normalizations, Weighting Matrices Iterated to Convergence**

| | a-Normalization | | | | ?-Normalization | | | |
|---|---|---|---|---|---|---|---|---|
| | $b^*$ | $\lambda_f^*$ | $R^2$ | $J$ | $b^\diamond$ | $\lambda_f^\diamond$ | $R^2$ | $J$ |
| **Consumption growth** | 145.7 | 1.49 | 0.39 | 28.1 | 90.1 | 0.24 | -9.3 | 73.5 |
| | (10.9) | (0.40) | | (0.256) | (39.1) | (0.10) | | (0.000) |
| **Durables growth** | 90.5 | 5.35 | 0.27 | 23.0 | 37.2 | 0.13 | -18.9 | 66.8 |
| | (5.6) | (10.35) | | (0.521) | (31.7) | (0.12) | | (0.000) |
| ***Rm-Rf*** | 5.8 | 4.24 | -9.5 | 63.7 | 3.0 | 1.98 | -1.09 | 70.1 |
| | (0.8) | (0.6) | | (0.000) | (0.9) | (0.54) | | (0.000) |
| **Consumption Factors** | | | | | | | | |
| Consumption growth | 48.7 | 1.83 | 0.95 | 18.1 | 78.5 | 0.22 | -12.1 | 58.5 |
| | (13.6) | (1.84) | | (0.751) | (42.7) | (0.11) | | (0.000) |
| Durables growth | 64.8 | 3.02 | | | 48.2 | 0.19 | | |
| | (8.2) | (2.98) | | | (34.1) | (0.13) | | |
| **Yogo Factors** | | | | | | | | |
| Consumption growth | 38.8 | 1.39 | 0.97 | 18.2 | 20.4 | 0.08 | -1.56 | 48.5 |
| | (16.4) | (1.39) | | (0.691) | (44.3) | (0.11) | | (0.001) |
| Durables growth | 68.0 | 2.82 | | | 58.7 | 0.20 | | |
| | (8.6) | (2.22) | | | (31.5) | (0.12) | | |
| *Rm* | 0.4 | 2.9 | | | 2.9 | 1.78 | | |
| | (0.4) | (3.05) | | | (1.1) | (0.56) | | |
| **Fama-French Factors** | | | | | | | | |
| *Rm-Rf* | 5.5 | 2.81 | -0.96 | 46.5 | 4.5 | 1.74 | 0.42 | 53.4 |
| | (0.8) | (0.7) | | (0.002) | (1.1) | (0.63) | | (0.000) |
| *SMB* | -1.1 | 0.50 | | | -0.9 | 0.30 | | |
| | (1.1) | (0.35) | | | (1.4) | (0.35) | | |
| *HML* | 7.6 | 1.80 | | | 7.0 | 1.40 | | |
| | (1.0) | (0.62) | | | (1.3) | (0.51) | | |

*Note*: Quarterly data, 1949–2005. The reports GMM stimates of $b^*$, from the SDF $m_t = 1 - f_t'b^*$, obtained using the moment restriction $E(R_t^e m_t) = 0$ and GMM estimates of $\mu$ and $b^\diamond$, from the SDF $m_t = 1 - (f_t - \mu)'b^\diamond$, obtained using the moment restrictions $E(R_t^e m_t) = 0$, $E(f_t - \mu) = 0$. Both sets of estimates are based on iterating the weighting matrix to convergence. The variables $R_t^e$ and $f_t$ are defined in the note to Table 3. Estimates of the factor risk premium $\hat{\lambda}_f = \mathrm{S}_f \hat{b}^\diamond$ are also reported (in percent), where $\mathrm{S}_f$ is the sample covariance matrix of $f_t$. GMM-VARHAC standard errors are reported in parentheses for $b^*$, $\hat{\mu}$, $\hat{b}^\diamond$, $\hat{\lambda}_f^*$ and $\hat{\lambda}_f^\diamond$. The table reports the $R^2$ measure of fit between the sample mean of $R_t^e$ and the predicted mean returns. For the $a$-normalization the predicted mean returns are $D_T\hat{b}^*$, where $D_T = \frac{1}{T}\sum_{t=1}^T R_t^e f_t'$. For the $\xi$-normalization they are $d_T\hat{b}^\diamond$, where $d_T = \frac{1}{T}\sum_{t=1}^T R_t^e(f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, $J$, is asymptotically distributed as a $\chi^2_{8-k}$, where $k$ is the number of risk factors. The p-value is in parentheses.