Variable Selection for Portfolio Choice*

Unpublished Technical Appendix

Yacine Aït-Sahalia  Michael W. Brandt
Department of Economics The Wharton School
Princeton University† University of Pennsylvania‡
and NBER and NBER

A Estimator and Distribution Theory

We let \( \alpha_0(Z_t) \) denote the solution to the unconstrained problem \( \max_{\alpha} \mathbb{E} [v(\alpha' R_{t+1}) | Z_t] \) and \( \alpha_0(Z_t' \beta; \beta) \) denote the solution to the constrained problem \( \max_{\alpha} \mathbb{E} [v(\alpha' R_{t+1}) | Z_t' \beta]. \) For the optimal portfolio weights \( \alpha_t \) to depend on the predictors \( Z_t \) only through the index \( Z_t' \beta, \) we assume that there exist a parameter vector \( \beta_0 \) such that:

\[
\mathbb{E} [v(\alpha_0(Z_t)' R_{t+1}) | Z_t] = \mathbb{E} [v(\alpha_0(Z_t' \beta_0)' R_{t+1}) | Z_t' \beta_0]
\]  

(A.1)

and \( \alpha_0(Z_t) = \alpha_0(Z_t' \beta_0; \beta_0) \) for almost all \( Z_t. \)

We assume that the data \( \{Z_t, R_{t+1}\}_{t=1}^T \) are stationary and strongly mixing at an exponential rate of decay. Then, we can obtain estimates of \( \hat{\alpha}(\cdot) \) and \( \hat{\beta} \) by solving the sample moments problem:

\[
\frac{1}{T} \sum_{t=1}^{T} m_{t+1}(\hat{\alpha}(Z_t' \hat{\beta}; \hat{\beta})) g(Z_t) = 0
\]  

(A.2)

\[
\frac{1}{T h} \sum_{t=1}^{T} m_{t+1}(\hat{\alpha}(Z_t' \hat{\beta}; \hat{\beta})) K \left( \frac{Z_t' \hat{\beta} - Z_t' \hat{\beta}}{h} \right) = 0 \quad \forall Z \in \{Z_t\}_{t=1}^T
\]  

(A.3)

where

\[
m_{t+1}(\alpha) = \frac{\partial v(\alpha' R_{t+1})}{\partial W} R_{t+1}
\]  

(A.4)

and \( g(\cdot) \) is a vector of instruments. Equation (A.2) identifies the index coefficients for a given portfolio policy, as described in Section II.B. Equation (A.3) characterizes the optimal portfolio policy for a given index composition through the first-order conditions of the portfolio optimization, as suggested

---

† Princeton, NJ 08544-1021. Phone: (609) 258-4015. E-mail: yacine@princeton.edu.
‡ Philadelphia, PA 19104-6367. Phone: (215) 898-3609. E-mail: brandtm@wharton.upenn.edu.
by Brandt (1999), where we replace the conditional expectation with a nonparametric regression. The function $K(\cdot)$ denotes the kernel of the nonparametric regression and $h$ is the kernel bandwidth. We choose the bandwidth as $O(T^{-1/5}\ln T)$, a rate just fast enough to asymptotically eliminate the bias of $\hat{\beta}$ induced by the nonparametric regression (at the cost of a slight increase in asymptotic variance, relative to the optimal mean-squared error tradeoff).

The asymptotics of $\hat{\beta}$ follow from a Taylor-expansion of the moments (A.2) and (A.3) around the true $\alpha_0$ and $\beta_0$. Starting with the first moment condition and expanding around $\beta_0$, we have:

$$
\frac{1}{T} \sum_{t=1}^{T} m_{t+1}(\hat{\alpha}(Z'_t \beta; \beta)) g(Z_t) +
$$

$$
\frac{1}{T} \sum_{t=1}^{T} \frac{\partial m_{t+1}(\hat{\alpha}(Z'_t \beta^*; \beta^*))}{\partial \alpha} \frac{\partial \hat{\alpha}(Z'_t \beta^*; \beta^*)}{\partial \beta} g(Z_t) \{\hat{\beta} - \beta_0\} = 0,
$$

(A.5)

where $\beta^* \in (\beta_0, \hat{\beta})$ with the usual convention for vector intervals. Rearranging this expression:

$$
\hat{\beta} - \beta_0 = - \left[ \frac{1}{T} \sum_{t=1}^{T} \frac{\partial m_{t+1}(\hat{\alpha}(Z'_t \beta^*; \beta^*))}{\partial \alpha} \frac{\partial \hat{\alpha}(Z'_t \beta^*; \beta^*)}{\partial \beta} g(Z_t) \right]^{-1} \times \left[ \frac{1}{T} \sum_{t=1}^{T} m_{t+1}(\hat{\alpha}(Z'_t \beta_0; \beta_0)) g(Z_t) \right]
$$

(A.6)

and then expanding around $\alpha_0$, we get:

$$
\hat{\beta} - \beta_0 = - D^{-1} \left[ \frac{1}{T} \sum_{t=1}^{T} m_{t+1}(\alpha_0(Z'_t \beta_0; \beta_0)) g(Z_t) + \right.
$$

$$
\frac{1}{T} \sum_{t=1}^{T} \frac{\partial m_{t+1}(\alpha^*(Z'_t \beta_0; \beta_0))}{\partial \alpha} g(Z_t) \{\hat{\alpha}(Z'_t \beta_0; \beta_0) - \alpha_0(Z'_t \beta_0; \beta_0)\} \right] (1 + o_p(1)),
$$

(A.7)

where

$$
D = E \left[ \frac{\partial m_{t+1}(\alpha_0(Z'_t \beta_0; \beta_0))}{\partial \alpha} \frac{\partial \alpha_0(Z'_t \beta_0; \beta_0)}{\partial \beta} g(Z_t) \right]
$$

(A.8)

and $\alpha^* \in (\alpha_0, \hat{\alpha})$. Notice that $E[m_{t+1}(\alpha_0(Z'_t \beta_0; \beta_0)) g(Z_t)] = 0$, which leaves us with only the second summation in expansion (A.7) to study.

To obtain an expression for $\{\hat{\alpha}(Z'_t \beta_0; \beta_0) - \alpha_0(Z'_t \beta_0; \beta_0)\}$, we expand the second moment condition:

$$
\frac{1}{Th} \sum_{t=1}^{T} m_{t+1}(\hat{\alpha}(Z'_t \beta; \beta)) K \left( \frac{Z'_t \hat{\beta} - Z'_t \hat{\beta}}{h} \right) +
$$

$$
\frac{1}{Th} \sum_{t=1}^{T} \frac{\partial m_{t+1}(\hat{\alpha}(Z'_t \beta^*; \beta^*))}{\partial \alpha} K \left( \frac{Z'_t \hat{\beta} - Z'_t \hat{\beta}}{h} \right) \{\hat{\alpha}(Z'_t \beta; \beta) - \alpha_0(Z'_t \beta; \hat{\beta})\} = 0
$$

(A.9)
which implies that for each $Z \in \{Z_t\}_{t=1}^T$:

$$
\hat{\alpha}(Z'_t; \hat{\beta}) - \alpha_0(Z'_t; \beta) = \left[ \frac{1}{T \hbar} \sum_{s=1}^{T} \frac{\partial m_{s+1}(\hat{\alpha}(Z'_s; \beta))}{\partial \alpha} K\left( \frac{Z'_s - Z'_t}{h} \right) \right]^{-1} \times \left[ \frac{1}{T \hbar} \sum_{s=1}^{T} m_{s+1}(\hat{\alpha}(Z'_s; \beta)) K\left( \frac{Z'_s - Z'_t}{h} \right) \right]
$$

\begin{equation}
= -A^{-1} \left[ \frac{1}{T \hbar} \sum_{s=1}^{T} m_{s+1}(\hat{\alpha}(Z'_s; \beta)) K\left( \frac{Z'_s - Z'_t}{h} \right) \right] (1 + o_p(1)) \tag{A.10}
\end{equation}

where

$$
A = E \left[ \frac{\partial m_{s+1}(\alpha(Z'_s; \beta_0))}{\partial \alpha} \right] \mathbf{1} (Z'_s = Z'_t). \tag{A.12}
$$

Evaluating expansion (A.11) at $Z = Z_t$ and substituting it into the second summation in equation (A.7) yields:

\begin{equation}
\frac{1}{T} \sum_{t=1}^{T} \frac{\partial m_{t+1}(\alpha(Z'_t; \beta_0))}{\partial \alpha} g(Z_t) \{ \hat{\alpha}(Z'_t; \beta_0) - \alpha_0(Z'_t; \beta_0) \} = \tag{A.13}
\end{equation}

\begin{equation}
- \frac{1}{T^2 \hbar} \sum_{t=1}^{T} \sum_{s=1}^{T} \frac{\partial m_{s+1}(\alpha(Z'_s; \beta_0))}{\partial \alpha} g(Z_t) A^{-1} m_{s+1}(\alpha(Z'_s; \beta_0)) K\left( \frac{Z'_s - Z'_t}{h} \right) \tag{A.14}
\end{equation}

Define $q_{t+1,s+1}(Z_t, Z_s)$ as:

\begin{equation}
\frac{\partial m_{s+1}(\alpha_0(Z'_s; \beta_0))}{\partial \alpha} g(Z_t) A^{-1} m_{s+1}(\alpha_0(Z'_s; \beta_0)) \frac{1}{h} K\left( \frac{Z'_s - Z'_t}{h} \right). \tag{A.14}
\end{equation}

The function $q_{t+1,s+1}(Z_t, Z_s)$ is not symmetric in $t$ and $s$. In order to write it in $U$-statistic form, we work with the symmetrized version:

\begin{equation}
p_{t+1,s+1}(Z_t, Z_s) = \frac{q_{t+1,s+1}(Z_t, Z_s) + q_{s+1,t+1}(Z_s, Z_t)}{2} \tag{A.15}
\end{equation}

and sum only on $s > t$. This allows us to rewrite equation (A.13) as:

\begin{equation}
\frac{1}{T} \sum_{t=1}^{T} \frac{\partial m_{t+1}(\alpha(Z'_t; \beta_0))}{\partial \alpha} g(Z_t) \{ \hat{\alpha}(Z'_t; \beta_0) - \alpha_0(Z'_t; \beta_0) \} = \tag{A.16}
\end{equation}

\begin{equation}
\left[ \frac{2}{T^2} \sum_{t=1}^{T} \sum_{s=t+1}^{T} p_{t+1,s+1}(Z_t, Z_s) \right] (1 + o_p(1))
\end{equation}
Let \( r(Z_t) = E[p_{t+1,s+1}(Z_t, Z_s) | Z_t] \). The two components of this expectation are:

\[
E[q_{t+1,s+1}(Z_t, Z_s)] = a_{t+1}(Z_t^0; \beta_0) E\left[b_{s+1}(Z_s^0; \beta_0) \frac{1}{h} K\left(\frac{Z_s^0 - Z_t^0}{h}\right) | Z_t\right]
\]

(A.17)

and

\[
E[q_{s+1,t+1}(Z_s, Z_t)] = E\left[a_{s+1}(Z_s^0; \beta_0) \frac{1}{h} K\left(\frac{Z_s^0 - Z_t^0}{h}\right) | Z_t\right] b_{t+1}(Z_t^0; \beta_0)
\]

(A.18)

Finally, the expansion (A.7) simplifies to:

\[
\sqrt{T}(\hat{\beta} - \beta_0) = -D^{-1}\left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} e_{t+1}(Z_t)\right] + o_p(1)
\]

(A.19)

where \( e_{t+1}(Z_t) = e_{t+1}^{(1)}(Z_t) + e_{t+1}^{(2)}(Z_t) + e_{t+1}^{(3)}(Z_t) \), with:

\[
e_{t+1}^{(1)}(Z_t) = m_{t+1}(\alpha_0(Z_t^0; \beta_0)) g(Z_t)
\]

(A.20)

\[
e_{t+1}^{(2)}(Z_t) = a_{t+1}(Z_t^0; \beta_0) E\left[b_{s+1}(Z_s^0; \beta_0) | Z_t\right] b_{t+1}(Z_t^0; \beta_0)
\]

Notice that \( E[e_{t+1}(Z_t)] = 0 \) and that:

\[
\frac{1}{\sqrt{T}} \sum_{t=1}^{T} e_{t+1}(Z_t) \overset{d}{\rightarrow} N[0, \Sigma],
\]

(A.21)

where \( \Sigma = \text{Var}[e_{t+1}(Z_t)] \) is the sum of the three variances and twice the sum of the three covariances of the terms \( e_{t}^{(i)} \), for \( i = 1, 2, 3 \), in equation (A.20). The final result is then:

\[
\sqrt{T}(\hat{\beta} - \beta_0) \overset{d}{\rightarrow} N[0, D^{-1}\Sigma D^{-1}],
\]

(A.22)