Internet Appendix to "Market Segmentation and Cross-Predictability of Returns"*

Lior Menzly
Nomura Asset Management

Oguzhan Ozbas
University of Southern California

This document contains supplementary material to the paper titled "Market Segmentation and Cross-Predictability of Returns." The document contains two sections. Section I studies a limited-information model whose predictions about cross-predictability are tested in the paper. Section II reports tables prepared in response to questions raised during the review process that may be of general interest to the reader, but are not reported in the paper.

*Citation format: Menzly, Lior, and Oguzhan Ozbas, 2010, Internet Appendix to "Market Segmentation and Cross-Predictability of Returns," *Journal of Finance* 65, 1555-1580, http://www.afajof.org/supplements.asp.

I. A Model of Cross-Predictability

In this section, we study a limited-information model in which dispersed information diffuses slowly across markets with correlated fundamentals and leads to cross-predictability in returns. The model is inspired by Hong and Stein (1999) with respect to dispersed information and by Hong, Torous, and Valkanov (2007) with respect to the study of markets with correlated fundamentals, and formally extends the latter in two directions: (i) we introduce uninformed investors (investors who do not have informative signals) to study their effect on cross-predictability, and (ii) we relax the assumption that informed investors invest only in the market about which they acquire informative signals to study the joint behavior of stock returns and informed trade across related markets.

The analysis proceeds in two steps. We first consider a single asset market in isolation to study return predictability. We then consider two asset markets with correlated fundamentals to study return cross-predictability.

A. Return Predictability in a Single Market

Suppose that there are three dates $\{t-1, t, t+1\}$, a single risky asset in zero supply that pays a liquidating dividend d at date $t+1$, and a riskless asset whose gross payoff is normalized to one and hence is the numeraire. (The zero-supply assumption is for simplicity and without loss of generality. A positive supply of the risky asset would merely lead to unconditional risk premia at dates $t-1$ and t, and hence would not affect the analysis.) There are n investors in the economy with constant absolute risk aversion parameter a. Investors trade the risky asset at dates $t-1$ and t with market clearing prices denoted p_{t-1} and p_t, respectively, and then consume the liquidating dividend at date $t+1$. Their common prior belief at date $t-1$ is that $d \sim N\left(\bar{d}, \sigma_d^2\right)$. At date t, an informative but noisy signal s about d arrives, where $s = d + \varepsilon$ and ε is an independent normally distributed noise term with mean zero and variance σ_ε^2. The informative signal allows investors who receive it to update their beliefs about d and adjust their demands for the risky asset at date t.

PROPOSITION 1 *When every investor receives the informative signal s about d, equilibrium prices do not exhibit predictability.*

Proof: After receiving the informative signal s at date t, investors solve the following optimization problem:

$$\max_{x_t} E\left[-e^{-aW_{t+1}} \mid s\right]. \tag{IA.1}$$

Substituting in $W_{t+1} = W_t - p_t x_t + d x_t$ and then evaluating the expectation, the optimization problem is

$$\max_{x_t} \; - e^{-a\left(W_t - p_t x_t + E_{d|s} x_t - \frac{1}{2} a \sigma_{d|s}^2 x_t^2\right)}. \tag{IA.2}$$

Investor demand for the risky asset at date t is therefore given by

$$x_t = \frac{E_{d|s} - p_t}{a\sigma_{d|s}^2}, \tag{IA.3}$$

where

$$E_{d|s} = \bar{d} + \underbrace{\frac{\sigma_d^2}{\sigma_d^2 + \sigma_\varepsilon^2}}_{\beta_s}\left(s - \bar{d}\right) \tag{IA.4}$$

$$\sigma_{d|s}^2 = \sigma_d^2\left(1 - \frac{\sigma_d^2}{\sigma_d^2 + \sigma_\varepsilon^2}\right). \tag{IA.5}$$

Posterior beliefs about the liquidating dividend come from a normal projection of s on d,

$$d = \bar{d} + \beta_s\left(s - \bar{d}\right) + \eta_s, \tag{IA.6}$$

where the residual uncertainty about the liquidating dividend η_s is distributed $N\left(0, \sigma_{d|s}^2\right)$. By the optimality of the projection,

$$\eta_s \perp \left(s - \bar{d}\right). \tag{IA.7}$$

Given that market clearing at date t requires $nx_t = 0$, substituting in investor demand yields

$$p_t = \bar{d} + \beta_s \left(s - \bar{d}\right). \tag{IA.8}$$

Note that p_t fully incorporates the informative signal s as given by the optimal projection.

Back at date $t-1$, investors solve the following optimization problem:

$$\max_{x_{t-1}} E\left[-e^{-aW_{t+1}}\right]. \tag{IA.9}$$

Substituting in $W_{t+1} = W_t - p_t(s)\,x_t(s) + dx_t(s)$ and $W_t = W_{t-1} - p_{t-1}x_{t-1} + p_t(s)\,x_{t-1}$, the optimization problem is

$$\max_{x_{t-1}} E\left[-e^{-a(W_{t-1}-p_{t-1}x_{t-1}+p_t(s)x_{t-1}-p_t(s)x_t(s)+dx_t(s))}\right]. \tag{IA.10}$$

Given that $p_t(s) = \bar{d} + \beta_s\left(s - \bar{d}\right)$ and $x_t(s) = 0$, we can write the optimization problem as

$$\max_{x_{t-1}} \; - e^{-a\left(W_{t-1}-\left(p_{t-1}-\bar{d}\right)x_{t-1}-\frac{1}{2}a\beta_s^2\sigma_s^2x_{t-1}^2\right)}. \tag{IA.11}$$

Investor demand for the risky asset at date $t-1$ is therefore given by

$$x_{t-1} = \frac{\bar{d} - p_{t-1}}{a\beta_s^2\sigma_s^2}. \tag{IA.12}$$

Given that market clearing at date $t-1$ requires $nx_{t-1} = 0$, substituting in investor demand yields

$$p_{t-1} = \bar{d}. \tag{IA.13}$$

Without loss of generality, define returns

$$r_t = p_t - p_{t-1} \tag{IA.14}$$

$$r_{t+1} = p_{t+1} - p_t. \tag{IA.15}$$

Evaluating the lagged beta of r_t on r_{t+1},

$$\frac{Cov\left(r_{t+1}, r_t\right)}{Var\left(r_t\right)} = \frac{Cov\left(d - \bar{d} - \beta_s\left(s - \bar{d}\right), \beta_s\left(s - \bar{d}\right)\right)}{Var\left(\beta_s\left(s - \bar{d}\right)\right)} \tag{IA.16}$$

$$= \frac{Cov\left(\eta_s, \beta_s\left(s - \bar{d}\right)\right)}{Var\left(\beta_s\left(s - \bar{d}\right)\right)} \tag{IA.17}$$

$$= 0 \qquad \left[\eta_s \perp \left(s - \bar{d}\right)\right], \tag{IA.18}$$

from which it is clear that equilibrium prices do not exhibit continuation because the informative signal is fully incorporated at $t = 1$.■

This neoclassical result follows from the fact that every investor adjusts his or her demand for the risky asset at date t after receiving s. When every individual demand incorporates the information in s, aggregate demand and p_t do so as well and equilibrium prices do not exhibit predictability – in the sense that the residual uncertainty $(d - p_t)$ left at date t is orthogonal to p_t. As the next proposition shows, however, when only a fraction $\alpha \in (0, 1)$ of the investor population receives the informative signal, and as a result investors differ in their information sets, equilibrium prices can exhibit predictability, in particular, continuation defined as $Cov\left(d - p_t, p_t - p_{t-1}\right) > 0$.

PROPOSITION 2 *When only a fraction $\alpha \in (0, 1)$ of the investor population receives the informative signal s about d, equilibrium prices exhibit continuation.*

Proof: For fraction α of the population (n_i/n), demand for the risky asset after receiving the informative signal s at date t is

$$x_t^i = \frac{E_{d|s} - p_t}{a\sigma_{d|s}^2}, \tag{IA.19}$$

whereas for fraction $(1 - \alpha)$ of the population (n_u/n), demand for the risky asset at date t is

$$x_t^u = \frac{\bar{d} - p_t}{a\sigma_d^2}. \tag{IA.20}$$

Given that market clearing at date t requires $\alpha n x_t^i + (1 - \alpha)\, n x_t^u = 0$, substituting in investor demands x_t^i and x_t^u yields

$$p_t = \bar{d} + \underbrace{\frac{\alpha\sigma_d^2}{\alpha\sigma_d^2 + (1 - \alpha)\,\sigma_{d|s}^2}}_{\gamma}\beta_s\left(s - \bar{d}\right). \tag{IA.21}$$

Note that $0 < \gamma < 1$ and hence p_t does not incorporate the informative signal s fully as given by the optimal projection.

Back at date $t-1$, investors solve the following optimization problem:

$$\max_{x_{t-1}} E\left[-e^{-aW_{t+1}}\right]. \tag{IA.22}$$

Substituting in $W_{t+1} = W_t - p_t(s)\,x_t(s) + dx_t(s)$, $W_t = W_{t-1} - p_{t-1}x_{t-1} + p_t(s)\,x_{t-1}$, and $p_t(s)$, the optimization problem is

$$\max_{x_{t-1}} E\left[-e^{-a\left(W_{t-1}-p_{t-1}x_{t-1}+\left(\bar{d}+\gamma\beta_s\left(s-\bar{d}\right)\right)x_{t-1}-\left(\bar{d}+\gamma\beta_s\left(s-\bar{d}\right)\right)x_t(s)+dx_t(s)\right)}\right]. \tag{IA.23}$$

Further substituting in $x_t^i(s) = \frac{E_{d|s}-p_t(s)}{a\sigma^2_{d|s}}$ for informed investors who will receive s at date t and taking the expectation yields

$$\max_{x^i_{t-1}} \; -C^i_{t-1} e^{-a\left(W_{t-1}-\left(p_{t-1}-\bar{d}\right)x_{t-1}-\frac{1}{2}a\sigma_d^2\left(\frac{\alpha^2\sigma_d^2}{\alpha^2\sigma_d^2+\sigma_\varepsilon^2}\right)x^2_{t-1}\right)}, \tag{IA.24}$$

where

$$C^i_{t-1} = \sqrt{\frac{\left(\alpha\sigma_d^2+\sigma_\varepsilon^2\right)^2}{\left(\sigma_d^2+\sigma_\varepsilon^2\right)\left(\alpha^2\sigma_d^2+\sigma_\varepsilon^2\right)}}. \tag{IA.25}$$

In computing the expectation, we use the result

$$E\left[-e^{-a\left(\xi+\psi\left(d-\bar{d}\right)+\phi\left(d-\bar{d}\right)^2\right)}\right] = -\frac{1}{\sqrt{1+2a\phi\sigma_d^2}} e^{-a\left(\frac{\xi\left(1+2a\phi\sigma_d^2\right)-\frac{1}{2}a\psi^2\sigma_d^2}{1+2a\phi\sigma_d^2}\right)} \tag{IA.26}$$

and the fact that $s = d+\varepsilon$ and ε is orthogonal to d. Solving for x^i_{t-1} yields

$$x^i_{t-1} = \frac{\left(\bar{d}-p_{t-1}\right)}{a\sigma_d^2}\frac{\alpha^2\sigma_d^2+\sigma_\varepsilon^2}{\alpha^2\sigma_d^2}. \tag{IA.27}$$

For uninformed investors who will not become informed at date t, substituting in $x_t^u = \frac{\bar{d}-p_t(s)}{a\sigma_d^2}$ yields

$$\max_{x^u_{t-1}} \; -C^u_{t-1} e^{-a\left(W_{t-1}-\left(p_{t-1}-\bar{d}\right)x_{t-1}-\frac{1}{2}a\sigma_d^2\left(\frac{\alpha^2\sigma_d^4+\alpha^2\sigma_d^2\sigma_\varepsilon^2}{\alpha^2\sigma_d^4+\alpha^2\sigma_d^2\sigma_\varepsilon^2+\sigma_\varepsilon^4}\right)x^2_{t-1}\right)}, \tag{IA.28}$$

where

$$C^u_{t-1} = \sqrt{\frac{\left(\alpha\sigma_d^2+\sigma_\varepsilon^2\right)^2\left(\alpha^2\sigma_d^4+\sigma_\varepsilon^4\right)}{\left(\alpha^2\sigma_d^4+2\sigma_d^2\sigma_\varepsilon^2+\sigma_\varepsilon^4\right)\left(\alpha^2\sigma_d^4+\alpha^2\sigma_d^2\sigma_\varepsilon^2+\sigma_\varepsilon^4\right)}}. \tag{IA.29}$$

Solving for x^u_{t-1} yields

$$x^u_{t-1} = \frac{\left(\bar{d} - p_{t-1}\right)}{a\sigma_d^2} \frac{\alpha^2\sigma_d^4 + \alpha^2\sigma_d^2\sigma_\varepsilon^2 + \sigma_\varepsilon^4}{\alpha^2\sigma_d^4 + \alpha^2\sigma_d^2\sigma_\varepsilon^2}. \tag{IA.30}$$

Given that market clearing at date $t-1$ requires $\alpha n x^i_{t-1} + (1-\alpha) n x^u_{t-1} = 0$, substituting in investor demands x^i_{t-1} and x^u_{t-1} yields

$$p_{t-1} = \bar{d}. \tag{IA.31}$$

Evaluating the lagged beta of r_t on r_{t+1},

$$\begin{aligned}
\frac{Cov\left(r_{t+1}, r_t\right)}{Var\left(r_t\right)} &= \frac{Cov\left(p_{t+1} - p_t, p_t - p_{t-1}\right)}{Var\left(p_t - p_{t-1}\right)} && \text{(IA.32)} \\
&= \frac{Cov\left(\left(1-\gamma\right)\beta_s\left(s - \bar{d}\right) + \eta_s, \gamma\beta_s\left(s - \bar{d}\right)\right)}{Var\left(\gamma\beta_s\left(s - \bar{d}\right)\right)} && \text{(IA.33)} \\
&= \frac{\left(1-\gamma\right)\gamma\beta_s^2\left(\sigma_d^2 + \sigma_\varepsilon^2\right)}{\gamma^2\beta_s^2\left(\sigma_d^2 + \sigma_\varepsilon^2\right)} && \text{(IA.34)} \\
&= \frac{1-\alpha}{\alpha}\frac{\sigma_\varepsilon^2}{\sigma_d^2 + \sigma_\varepsilon^2}, && \text{(IA.35)}
\end{aligned}$$

equilibrium prices exhibit continuation because the informative signal is not fully incorporated at date t.■

Equilibrium prices exhibit continuation because some investors do not receive s and also fail to infer s from publicly available information p_t to adjust their demand for the risky asset at date t. While informed investors adjust their demand, due to limited risk-bearing capacity they do not completely make up for the lack of adjustment in uninformed demand. As a result, aggregate demand and p_t incorporate the information in s only partly and equilibrium prices exhibit continuation – in the sense that the residual uncertainty $(d - p_t)$ left at date t is positively correlated with p_t. This feature of the model is common to a broad class of "disagreement models" as articulated by Hong and Stein (2007). Skill-based differences in information acquisition and processing costs among investors could plausibly result in heterogeneous beliefs and lead to equilibria in which investors with information acquisition and processing costs below a certain threshold choose to become informed and others choose to remain uninformed. Moreover, the magnitude of continuation decreases in

α. This is because the more informed investors there are in the market, the more information is impounded into p_t, and the less predictable residual uncertainty is left at date t.

B. Cross-Predictability Between Two Markets

We now turn to return cross-predictability. Suppose that there are two risky assets $k \in \{1, 2\}$ both in zero supply paying correlated liquidating dividends d_1 and d_2 at date $t+1$. The common prior belief at date $t-1$ is that $(d_1, d_2) \sim N\left(\bar{d}1, \Sigma\right)$, where

$$\Sigma = \begin{bmatrix} \sigma_d^2 & \rho\sigma_d^2 \\ \rho\sigma_d^2 & \sigma_d^2 \end{bmatrix}. \tag{IA.36}$$

At date t, two informative but noisy signals, s_1 about d_1 and s_2 about d_2, arrive, where $s_1 = d_1 + \varepsilon_1$, $s_2 = d_2 + \varepsilon_2$, and ε_1 and ε_2 are independent normally distributed noise terms with mean zero and variance σ_ε^2. Reflecting the specialization of market participants in gathering information about only a subset of assets, one group of investors (fraction α_1 of the investor population) receives s_1 and another group of investors (fraction α_2 of the investor population) receives s_2. For simplicity, we assume that the two groups, which respectively receive informative signals s_1 and s_2, are disjoint.

PROPOSITION 3 *When fraction α_1 of the investor population receives the signal s_1 and another fraction α_2 of the investor population receives the signal s_2, equilibrium prices exhibit cross-predictability.*

Proof: For fraction α_k of the population, demand for the risky assets after receiving the informative signal s_k at date t for $k \in \{1, 2\}$ and $j \neq k$ is

$$X_t^{i[s_k]} = \frac{1}{a}\Sigma_{s_k}^{-1}\left(\begin{bmatrix} E_{d_1|s_k} \\ E_{d_2|s_k} \end{bmatrix} - P_t\right), \tag{IA.37}$$

where

$$E_{d_k|s_k} = \bar{d} + \frac{\sigma_d^2}{\sigma_d^2 + \sigma_\varepsilon^2}\left(s_k - \bar{d}\right) \tag{IA.38}$$

$$E_{d_j|s_k} = \bar{d} + \frac{\rho\sigma_d^2}{\sigma_d^2 + \sigma_\varepsilon^2}\left(s_k - \bar{d}\right) \tag{IA.39}$$

$$\sigma^2_{d_k|s_k} = \sigma_d^2\left(1 - \frac{\sigma_d^2}{\sigma_d^2 + \sigma_\varepsilon^2}\right) \tag{IA.40}$$

$$\sigma^2_{d_j|s_k} = \sigma_d^2\left(1 - \rho^2\frac{\sigma_d^2}{\sigma_d^2 + \sigma_\varepsilon^2}\right) \tag{IA.41}$$

$$\Sigma_{s_k} = \begin{bmatrix} \sigma^2_{d_1|s_k} & \rho\sigma_d^2\left(1 - \frac{\sigma_d^2}{\sigma_d^2+\sigma_\varepsilon^2}\right) \\ \rho\sigma_d^2\left(1 - \frac{\sigma_d^2}{\sigma_d^2+\sigma_\varepsilon^2}\right) & \sigma^2_{d_2|s_k} \end{bmatrix}. \tag{IA.42}$$

For fraction $(1 - \alpha_1 - \alpha_2)$ of the population, demand for the risky asset at date t is

$$X_t^u = \frac{1}{a}\Sigma^{-1}\left(\bar{d}1 - P_t\right). \tag{IA.43}$$

Given that market clearing at date t requires $\alpha_1 n X_t^{i[s_1]} + \alpha_2 n X_t^{i[s_2]} + (1 - \alpha_1 - \alpha_2)\, n X_t^u = 0$, substituting in investor demands $X_t^{i[s_1]}$, $X_t^{i[s_2]}$, and X_t^u for $k \in \{1, 2\}$ and $j \neq k$ yields

$$p_{k:t} = \bar{d} + \underbrace{\frac{\alpha_k\alpha_j\left(1-\rho^2\right)\sigma_d^4 + \alpha_k\sigma_d^2\sigma_\varepsilon^2}{\alpha_1\alpha_2\left(1-\rho^2\right)\sigma_d^4 + \left(\alpha_1+\alpha_2\right)\sigma_d^2\sigma_\varepsilon^2 + \sigma_\varepsilon^4}}_{\gamma_{d_k,s_k}}\left(s_k - \bar{d}\right) \tag{IA.44}$$

$$+ \underbrace{\frac{\alpha_j\rho\sigma_d^2\sigma_\varepsilon^2}{\alpha_1\alpha_2\left(1-\rho^2\right)\sigma_d^4 + \left(\alpha_1+\alpha_2\right)\sigma_d^2\sigma_\varepsilon^2 + \sigma_\varepsilon^4}}_{\gamma_{d_k,s_j}}\left(s_j - \bar{d}\right). \tag{IA.45}$$

As for equilibrium prices at date $t-1$, lengthy calculations that are similar to those in the proof of Proposition 2 yield

$$p_{1:t-1} = p_{2:t-1} = \bar{d}. \tag{IA.46}$$

For brevity, we omit these lengthy calculations and note that in any case $p_{1:t-1}$ and $p_{2:t-1}$ enter only as constants in the cross-predictability expressions below and therefore they are not key to establishing the claim of the proposition.

In addition, note that for $k \in \{1,2\}$ and $j \neq k$, the normal projection of s_k and s_j on d_k is given by

$$d_k = \bar{d} + \underbrace{\frac{\left(1-\rho^2\right)\sigma_d^4 + \sigma_d^2\sigma_\varepsilon^2}{\left(1-\rho^2\right)\sigma_d^4 + 2\sigma_d^2\sigma_\varepsilon^2 + \sigma_\varepsilon^4}}_{\beta_{d_k,s_k}}\left(s_k - \bar{d}\right) \tag{IA.47}$$

$$+\underbrace{\frac{\rho\sigma_d^2\sigma_\varepsilon^2}{\left(1-\rho^2\right)\sigma_d^4 + 2\sigma_d^2\sigma_\varepsilon^2 + \sigma_\varepsilon^4}}_{\beta_{d_k,s_j}}\left(s_j - \bar{d}\right) + \eta_{d_k}. \tag{IA.48}$$

By the optimality of the projections,

$$\eta_{d_1}, \eta_{d_2} \perp \left(s_1 - \bar{d}\right), \ \left(s_2 - \bar{d}\right). \tag{IA.49}$$

Without loss of generality, we define returns $r_{k:t} = p_{k:t} - p_{k:t-1}$ and $r_{k:t+1} = p_{k:t+1} - p_{k:t}$ for $k \in \{1,2\}$ as before and evaluate the lagged cross-beta of $r_{j:t}$ on $r_{k:t+1}$ for $j \neq k$, where

$$\begin{aligned} Cov\left(r_{k:t+1}, r_{j:t}\right) &= Cov(\left(\beta_{d_k,s_k} - \gamma_{d_k,s_k}\right)\left(s_k - \bar{d}\right) + \left(\beta_{d_k,s_j} - \gamma_{d_k,s_j}\right)\left(s_j - \bar{d}\right) + \eta_{d_k}, \\ &\quad \gamma_{d_j,s_k}\left(s_k - \bar{d}\right) + \gamma_{d_j,s_j}\left(s_j - \bar{d}\right)) \end{aligned} \tag{IA.50}$$

$$\begin{aligned} &= \left(\left(\beta_{d_k,s_k} - \gamma_{d_k,s_k}\right)\gamma_{d_j,s_k} + \left(\beta_{d_k,s_j} - \gamma_{d_k,s_j}\right)\gamma_{d_j,s_j}\right)\left(\sigma_d^2 + \sigma_\varepsilon^2\right) \\ &\quad + \left(\left(\beta_{d_k,s_k} - \gamma_{d_k,s_k}\right)\gamma_{d_j,s_j} + \left(\beta_{d_k,s_j} - \gamma_{d_k,s_j}\right)\gamma_{d_j,s_k}\right)\rho\sigma_d^2 \end{aligned} \tag{IA.51}$$

and

$$Var\left(r_{j:t}\right) = Var\left(\gamma_{d_j,s_k}\left(s_k - \bar{d}\right) + \gamma_{d_j,s_j}\left(s_j - \bar{d}\right)\right) \tag{IA.52}$$

$$= \left(\gamma_{d_j,s_k}^2 + \gamma_{d_j,s_j}^2\right)\left(\sigma_d^2 + \sigma_\varepsilon^2\right) + 2\gamma_{d_j,s_k}\gamma_{d_j,s_j}\rho\sigma_d^2. \tag{IA.53}$$

Further substituting in β_{d_k,s_k}, β_{d_k,s_j}, γ_{d_k,s_k}, γ_{d_k,s_j}, γ_{d_j,s_k}, and γ_{d_j,s_j}, the lagged cross-beta is

$$\rho\sigma_\varepsilon^4 \frac{\left(1-\rho^2\right)\alpha_k\alpha_j\left(2-\alpha_k-\alpha_j\right)\sigma_d^2 + \left(\alpha_k\left(1-\alpha_k\right) + \alpha_j\left(1-\alpha_j\right)\right)\sigma_\varepsilon^2}{\rho^2\alpha_k^2\sigma_\varepsilon^4\left(\sigma_d^2+\sigma_\varepsilon^2\right) + 2\rho^2\alpha_k\alpha_j\left(\left(1-\rho^2\right)\alpha_k\sigma_d^2 + \sigma_\varepsilon^2\right) + \alpha_j^2\left(\sigma_d^2+\sigma_\varepsilon^2\right)\left(\left(1-\rho^2\right)\alpha_k\sigma_d^2 + \sigma_\varepsilon^2\right)^2}, \tag{IA.54}$$

which shows that equilibrium prices exhibit cross-predictability in the sign of ρ when fundamental payoffs are correlated ($\rho \neq 0$).■

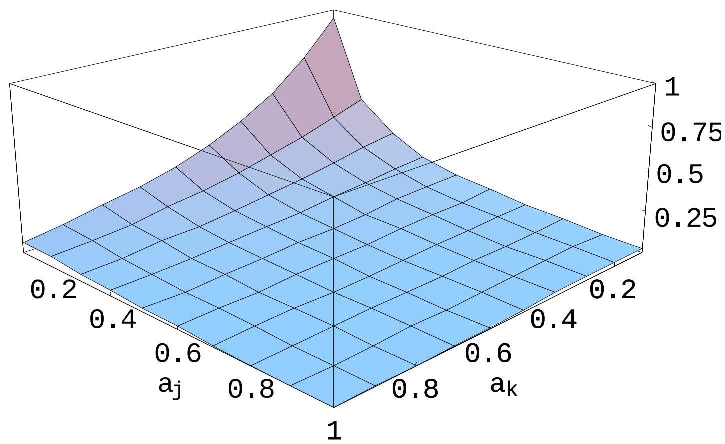

Figure IA.1. Lagged cross-beta of $r_{j:t}$ **on** $r_{k:t+1}$ ($\rho = 0.5$**,** $\sigma_d = 0.8$**,** $\sigma_\epsilon = 0.4$).

Equilibrium prices exhibit cross-predictability for the same reasons that they exhibit continuation. Some investors do not receive s_1 and hence do not adjust their demand for the first risky asset at date t. Likewise, some investors do not receive s_2 and hence do not adjust their demand for the second risky asset at date t. Consequently, both $p_{1:t}$ and $p_{2:t}$ incorporate the information in s_1 and s_2 only partly and equilibrium prices exhibit cross-predictability – in the sense that the residual uncertainty left in $(d_k - p_{k:t})$ at date t is correlated with $p_{j:t}$ for $k \in \{1, 2\}$ and $j \neq k$.

C. Testable Predictions

In the model, the cross-predictability effect in returns declines with the number of informed investors in the market. This is because informative signals received by informed investors in the intermediate stage are incorporated into prices more fully when there are more informed investors. Figure IA.1 plots the relation between the presence of informed investors and cross-predictability.

The model also sheds light on how informed investors trade to exploit their informational advantage over uninformed investors. Specifically, when informed investors trade in one of the markets due to new information, they also trade in the other market. Previous work finds evidence in support of this pattern in the context of a single market – institutional investors trade to take advantage of the continuation effect in prices (Cohen, Gompers, and Vuolteenaho (2002)). Hence, an untested prediction of the model is whether institutional investors also trade to take advantage of the cross-predictability effect in prices.

II. Supplementary Results

A. Single-Segment vs. Multi-Segment Firms

In the paper, each stock is assigned to a BEA industry based on the stock's reported SIC or NAICS code in COMPUSTAT, which represents the firm's main business. While this is likely to be a good approximation for single-segment firms whose operations are concentrated in one industry, it is not clear whether this is also a good approximation for multi-segment firms that operate in multiple industries. To investigate this issue, we estimate the first specification in Table II for single-segment and multi-segment firms separately. To form these two samples, which are mutually exclusive, we use information from COMPUSTAT's segment files. If a firm is reported as having only one segment for the time period in question, we classify the firm as a single-segment firm. If the firm is instead reported as having more than one segment, we classify the firm as a multi-segment firm. In assigning multi-segment firms to BEA industries, we follow the same procedure as in the paper and use the reported SIC or NAICS code in COMPUSTAT, which has the desired property of representing the firm's main business and thus its main economic exposure.

The results of this exercise are reported in Table IA.I. The coefficient estimates for single-segment and multi-segment firms are presented in columns 1 and 2, respectively. The coefficient estimates in both columns are similar to those for the whole sample reported in column 1 in Table II, and are also similar to each other. The t-statistics are lower than before due to smaller sample sizes and a shorter sample period (COMPUSTAT's segment files only start in 1979). Thus, compositional issues and the potential industry misclassification of multi-segment firms do not appear to have a significant impact on our analyses.

B. *Differences in Expected Returns Across Analyst Coverage and Institutional Ownership Quintiles*

In addition to the specifications reported in Table III, we estimate additional specifications that allow for cross-sectional differences in expected returns across the different analyst coverage and institutional ownership quintiles. Specifically, Table IA.II reports panel regressions with monthly fixed effects and appropriate monthly clustering of standard errors, instead of Fama-MacBeth (1973) regressions, mainly to improve the efficiency of the estimates since the specification is significantly longer with direct quintile effects.

To provide a benchmark, Panel A reports estimates from specifications without the direct quintile effects. These estimates are similar to those reported in Table III. Panel B reports estimates from specifications with the direct quintile effects. Again, the primary coefficients of interest, namely, quintile interactions with lagged returns in related industries, are similar to those in Panel A and Table III.

C. *Small Stocks*

By excluding stocks with market capitalization below the 20th NYSE percentile, column 2 in Table II addresses the possibility that thin markets might be driving the stock-level cross-predictability results. Table IA.III repeats the same analysis for Table III. While the spreads between the low and high quintile interactions are smaller than those in Table III, the declining pattern of cross-predictability across the quintile interactions is still evident.

D. Difference in Data Frequency: Quarterly Institutional Ownership and Monthly Stock Returns

The Fama-MacBeth (1973) regressions in column 3 of Table III rely on quarterly institutional ownership data. For each monthly cross-sectional regression, we sort stocks into quintiles based on their level of institutional ownership in the previous quarter. This procedure implies that we use institutional ownership as of December of Year X-1 to sort stocks in January, February, and March of Year X, institutional ownership as of March of Year X to sort stocks in April, May, and June of Year X, institutional ownership as of June of Year X to sort stocks in July, August, and September of Year X, and institutional ownership as of September of Year X to sort stocks in October, November, and December of Year X.

A potential statistical issue with this procedure is that the use of the same quarterly institutional ownership data in three separate monthly cross-sectional regressions may induce correlation among the estimated coefficients, in which case the standard errors may be understated. Although this is unlikely to be a problem because the estimated coefficients are interactions of institutional ownership quintiles and lagged returns in related industries (which differ across monthly regressions), we investigate this concern by computing robust standard errors that account for the correlation of coefficient estimates within a given quarter. The results of this exercise are reported in Table IA.IV (corresponding to column 3 in Table III). The t-statistics with robust standard errors are only slightly smaller, and none of the conclusions is affected.

E. Trading Strategies Excluding Small Stocks, and Alternative Trading Strategies

To address the general concern that trading profits may be driven by small stocks, the paper considers trading strategies that buy and sell value-weighted industry portfolios in Table V. To further address the concern that value-weighting may not be enough (because low capitalization

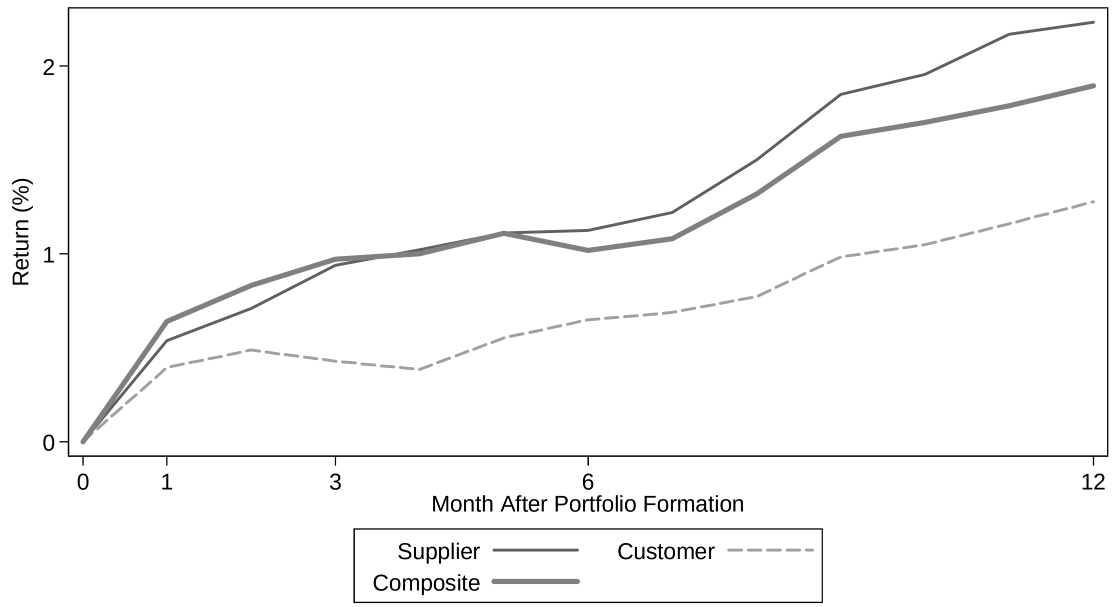

Figure IA.2. Performance of trading strategies in event time.

stocks still need to be bought and sold), we repeat the analysis in Table V by excluding stocks with market capitalization below the 20th NYSE percentile. Table IA.V presents the results of this analysis, and shows that the trading profits reported in the paper are not driven by small stocks.

A related analysis in Table IA.VI explores different formation and holding periods, where there are low-volume trading strategies with holding periods as long as 12 months that yield more than 2%. Finally, Figure IA.2 shows the performance of trading strategies in event time.

F. BEA Surveys

The Use Table data on the inter-industry flow of goods and services that we use to identify supplier and customer industries (see Section II.A.2, Benchmark Input-Output Surveys) are freely available from the Bureau of Economic Analysis and can be downloaded from their web site (http://www.bea.gov/industry/index.htm#benchmark_io, accessed on June 5, 2009). Table

IA.VII lists the industries in the 1987 survey, and Table IA.VIII provides the dictionary linking SIC codes to industries.

References

Cohen, Randolph, Paul Gompers, and Tuomo Vuolteenaho, 2002, Who underreacts to cash-flow news? Evidence from trading between individuals and institutions, *Journal of Financial Economics* 66, 409-462.

Fama, Eugene, and James MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

Hong, Harrison, and Jeremy Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143-2184.

Hong, Harrison, and Jeremy Stein, 2007, Disagreement and the stock market, *Journal of Economic Perspectives* 21, 109-128.

Hong, Harrison, Walter Torous, and Rossen Valkanov, 2007, Do industries lead stock markets? *Journal of Financial Economics* 83, 367-396.

Table IA.I

Cross-Predictability Effects for Single-Segment and Multi-Segment Firms

This table presents time-series averages of coefficient estimates from monthly cross-sectional regressions of stock returns. The sample includes single-segment firms in column 1, and multi-segment firms in column 2. Supplier (customer) returns consist of supplier (customer) industry returns weighted by the inter-industry flow of goods and services reported in the Benchmark Input-Output Surveys of the Bureau of Economic Analysis. All return variables are in excess of the risk-free rate. *t*-statistics are reported in parentheses. Standard errors assume independence across monthly regressions. ***, **, or * indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

	(1)	(2)
Constant	0.006	0.008**
	(1.56)	(2.58)
$r_{supplier,t-1}$	0.102***	0.111***
	(2.96)	(3.84)
$r_{customer,t-1}$	0.078***	0.074***
	(2.84)	(3.52)
$r_{stock,t-1}$	-0.061***	-0.062***
	(12.23)	(11.07)
$r_{stock,t-2:t-12}$	0.004***	0.003**
	(2.83)	(2.07)
$r_{industry,t-1}$	0.133***	0.122***
	(9.75)	(11.82)
R^2	0.025	0.028
T	318	318
Sample:	Single-segment	Multi-segment

Table IA.II

Analyst Coverage, Institutional Ownership, and Cross-Predictability Effects

This table presents panel regressions in which monthly stock returns are regressed on lagged related industry returns interacted with lagged analyst coverage and institutional ownership. $r_{composite}$ represents returns in related industries, and is calculated as the average of $r_{supplier}$ and $r_{customer}$. Analyst coverage for a stock in a given month is measured as the number of analysts who made an EPS forecast for the stock within the last 12 months (column 1) or the number of analysts who made an EPS forecast for the stock in that month (column 2). Institutional ownership is measured as the percentage of outstanding shares owned by institutions (column 3). Stocks are ranked into five quintiles based on analyst coverage and institutional ownership. All return variables are in excess of the risk-free rate. All specifications include year-month fixed effects. Robust standard errors (heteroskedasticity consistent and adjusted for clustering at the year-month level) are reported in brackets. ***, **, or * indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

Panel A: Without Own Effects			
	(1)	(2)	(3)
Constant	0.007***	0.007***	0.007***
	[0.001]	[0.001]	[0.001]
$r_{composite,t-1}$ *x Rank* $_{t-1}$ *(1st Quintile - Low)*	0.287**	0.277**	0.329***
	[0.120]	[0.119]	[0.106]
$r_{composite,t-1}$ *x Rank* $_{t-1}$ *(2nd Quintile)*	0.246**	0.217*	0.283***
	[0.119]	[0.122]	[0.109]
$r_{composite,t-1}$ *x Rank* $_{t-1}$ *(3rd Quintile)*	0.167	0.171	0.206*
	[0.122]	[0.120]	[0.107]
$r_{composite,t-1}$ *x Rank* $_{t-1}$ *(4th Quintile)*	0.083	0.107	0.128
	[0.126]	[0.125]	[0.111]
$r_{composite,t-1}$ *x Rank* $_{t-1}$ *(5th Quintile - High)*	-0.003	-0.002	0.048
	[0.135]	[0.131]	[0.115]
R^2	0.119	0.119	0.091
N obs	967,217	967,217	1,544,198

Panel B: With Own Effects			
	(1)	(2)	(3)
Constant	0.007***	0.004***	0.006***
	[0.001]	[0.001]	[0.002]
$Rank_{t-1}$ (2^{nd} Quintile)	-0.002*	0.002***	0.002
	[0.001]	[0.001]	[0.001]
$Rank_{t-1}$ (3^{rd} Quintile)	-0.001	0.003***	0.001
	[0.001]	[0.001]	[0.002]
$Rank_{t-1}$ (4^{th} Quintile)	0.000	0.004***	0.002
	[0.002]	[0.001]	[0.002]
$Rank_{t-1}$ (5^{th} Quintile - High)	0.001	0.004**	0.002
	[0.002]	[0.002]	[0.003]
$r_{composite,t-1}$ x $Rank_{t-1}$ (1^{st} Quintile - Low)	0.287**	0.283**	0.332***
	[0.121]	[0.119]	[0.107]
$r_{composite,t-1}$ x $Rank_{t-1}$ (2^{nd} Quintile)	0.250**	0.218*	0.282**
	[0.119]	[0.123]	[0.109]
$r_{composite,t-1}$ x $Rank_{t-1}$ (3^{rd} Quintile)	0.169	0.170	0.207*
	[0.122]	[0.120]	[0.107]
$r_{composite,t-1}$ x $Rank_{t-1}$ (4^{th} Quintile)	0.082	0.104	0.127
	[0.125]	[0.125]	[0.111]
$r_{composite,t-1}$ x $Rank_{t-1}$ (5^{th} Quintile - High)	-0.006	-0.008	0.046
	[0.135]	[0.131]	[0.115]
R^2	0.119	0.119	0.091
N obs	967,217	967,217	1,544,198

Table IA.III

Analyst Coverage, Institutional Ownership, and Cross-Predictability Effects

This table presents time-series averages of coefficient estimates from monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with lagged analyst coverage and institutional ownership. The sample excludes stocks with market capitalization below the 20th NYSE percentile. $r_{composite}$ represents returns in related industries, and is calculated as the average of $r_{supplier}$ and $r_{customer}$. Analyst coverage for a stock in a given month is measured as the number of analysts who made an EPS forecast for the stock within the last 12 months (column 1) or the number of analysts who made an EPS forecast for the stock in that month (column 2). Institutional ownership is measured as the percentage of outstanding shares owned by institutions (column 3). Stocks are ranked into five quintiles based on analyst coverage and institutional ownership. All return variables are in excess of the risk-free rate. *t*-statistics are reported in parentheses. Standard errors assume independence across monthly regressions. ***, **, or * indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

	(1)	(2)	(3)
Constant	0.009**	0.009**	0.010**
	(2.46)	(2.39)	(2.60)
$r_{composite,t-1}$ x $Rank_{t-1}$ (1st Quintile - Low)	0.250***	0.229***	0.242***
	(4.22)	(3.87)	(3.81)
$r_{composite,t-1}$ x $Rank_{t-1}$ (2nd Quintile)	0.223***	0.214***	0.213***
	(3.77)	(3.80)	(3.67)
$r_{composite,t-1}$ x $Rank_{t-1}$ (3rd Quintile)	0.187***	0.213***	0.202***
	(3.15)	(3.58)	(3.62)
$r_{composite,t-1}$ x $Rank_{t-1}$ (4th Quintile)	0.123**	0.137**	0.149***
	(2.18)	(2.29)	(2.73)
$r_{composite,t-1}$ x $Rank_{t-1}$ (5th Quintile - High)	0.100	0.096	0.110**
	(1.62)	(1.64)	(1.99)
R^2	0.018	0.017	0.017
T	281	281	303

Table IA.IV

Institutional Ownership and Cross-Predictability Effects

This table presents time-series averages of coefficient estimates from monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with lagged institutional ownership. $r_{composite}$ represents returns in related industries, and is calculated as the average of $r_{supplier}$ and $r_{customer}$. Institutional ownership is measured as the percentage of outstanding shares owned by institutions. Stocks are ranked into five quintiles based on institutional ownership. All return variables are in excess of the risk-free rate. *t*-statistics are reported in parentheses. Standard errors are heteroskedasticity consistent and adjusted for clustering at the year-quarter level. ***, **, or * indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

	(1)
Constant	0.008*
	(1.89)
$r_{composite,t-1}$ *x Rank* $_{t-1}$ *(1st Quintile - Low)*	0.380***
	(5.11)
$r_{composite,t-1}$ *x Rank* $_{t-1}$ *(2nd Quintile)*	0.317***
	(4.90)
$r_{composite,t-1}$ *x Rank* $_{t-1}$ *(3rd Quintile)*	0.244***
	(4.12)
$r_{composite,t-1}$ *x Rank* $_{t-1}$ *(4th Quintile)*	0.177***
	(2.96)
$r_{composite,t-1}$ *x Rank* $_{t-1}$ *(5th Quintile - High)*	0.067
	(1.12)
R^2	0.012
T	303

Table IA.V

Self-Financing Trading Strategies

This table reports the mean and standard deviation of monthly excess returns on value-weighted portfolios of industries formed on the basis of related industry returns in the previous month (reported figures are annualized). Stocks with market capitalization below the 20th NYSE percentile are excluded from industry portfolios. Industries are sorted into five bins at the beginning of each month according to returns in related industries in the previous month. Self-financing trading strategies reported in the last column consist of buying the high (5) portfolio (top quintile) and selling the low (1) portfolio (bottom quintile).

	Low (1)	(2)	(3)	(4)	High (5)	H - L
Panel A: Industries Sorted on $r_{supplier,t-1}$						
Mean return	0.027	0.053	0.051	0.087	0.093	0.066
Standard deviation	0.159	0.176	0.176	0.178	0.164	0.112
Sharpe ratio	0.169	0.302	0.292	0.490	0.564	0.587
Panel B: Industries Sorted on $r_{customer,t-1}$						
Mean return	0.018	0.052	0.059	0.067	0.083	0.065
Standard deviation	0.178	0.165	0.154	0.167	0.185	0.136
Sharpe ratio	0.099	0.316	0.380	0.400	0.448	0.480
Panel C: Industries Sorted on $r_{composite,t-1}$						
Mean return	0.012	0.041	0.063	0.073	0.090	0.078
Standard deviation	0.170	0.165	0.166	0.177	0.170	0.131
Sharpe ratio	0.072	0.246	0.381	0.413	0.531	0.592

Table IA.VI

Alternative Formation and Holding Periods

This table reports the monthly profitability of self-financing trading strategies formulated on the basis of lagged returns in related industries with various formation and holding periods (reported figures are annualized). For each trading strategy considered, industries are sorted at the beginning of each month into five bins according to their previous J-month related industry returns. The trading strategy then buys the high (5) portfolio (comprised of industries with previous J-month related industry returns in the top quintile), sells the low (1) portfolio (comprised of industries with previous J-month related industry returns in the bottom quintile), and holds the position for K months. As a result, the strategy holds in any given month a series of K portfolios that are selected in that month and as far back as K-1 months prior. *t*-statistics are reported in parentheses.

	K			
J	1	3	6	12
Panel A: Supplier Strategy				
1	0.073	0.035	0.016	0.017
	(4.27)	(3.23)	(1.91)	(2.70)
3	0.062	0.034	0.015	0.017
	(3.70)	(2.52)	(1.40)	(2.03)
6	0.032	0.015	0.017	0.015
	(1.91)	(1.03)	(1.27)	(1.38)
12	0.053	0.040	0.032	0.019
	(3.05)	(2.54)	(2.12)	(1.35)
Panel B: Customer Strategy				
1	0.070	0.021	0.016	0.016
	(3.37)	(1.60)	(1.53)	(2.07)
3	0.035	0.017	0.015	0.019
	(1.69)	(1.04)	(1.12)	(1.81)
6	0.037	0.033	0.037	0.028
	(1.74)	(1.68)	(2.10)	(1.87)
12	0.063	0.050	0.034	0.023
	(2.74)	(2.35)	(1.68)	(1.23)
Panel C: Composite Strategy				
1	0.087	0.041	0.020	0.021
	(4.26)	(3.23)	(2.04)	(2.89)
3	0.059	0.027	0.012	0.023
	(2.89)	(1.63)	(0.94)	(2.39)
6	0.052	0.026	0.026	0.027
	(2.39)	(1.41)	(1.64)	(2.06)
12	0.054	0.041	0.027	0.021
	(2.72)	(2.22)	(1.53)	(1.27)

Table IA.VII

BEA Industries

This table lists the industries in the 1987 Benchmark Input-Output Survey of the Bureau of Economic Analysis.

BEA Industry	Industry Name
1+2	Livestock and livestock products, and other agricultural products
3	Forestry and fishery products
4	Agricultural, forestry, and fishery services
5+6	Metallic ores mining
7	Coal mining
8	Crude petroleum and natural gas
9+10	Nonmetallic minerals mining
11+12	Construction
13	Ordnance and accessories
14	Food and kindred products
15	Tobacco products
16	Broad and narrow fabrics, yarn and thread mills
17	Miscellaneous textile goods and floor coverings
18	Apparel
19	Miscellaneous fabricated textile products
20+21	Lumber and wood products
22	Household furniture and fixtures
23	Non-household furniture and fixtures
24	Paper and allied products, except containers
25	Paperboard containers and boxes
26	Newspapers and periodicals, and other printing and publishing
27	Industrial and other chemicals, and agricultural fertilizers and chemicals
28	Plastics and synthetic materials
29	Drugs, and cleaning and toilet preparations
30	Paints and allied products
31	Petroleum refining and related products
32	Rubber and miscellaneous plastics products
33+34	Footwear, leather, and leather products
35	Glass and glass products
36	Stone and clay products
37	Primary iron and steel manufacturing
38	Primary nonferrous metals manufacturing
39	Metal containers
40	Heating, plumbing, and fabricated structural metal products
41	Screw machine products and stampings
42	Other fabricated metal products
43	Engines and turbines
44	Farm machinery
45	Construction and mining machinery
46	Materials handling machinery and equipment
47	Metalworking machinery and equipment
48	Special industry machinery and equipment

BEA Industry	Industry Name
49	General industrial machinery and equipment
50	Miscellaneous machinery, except electrical
51	Computer and office equipment
52	Service industry machinery
53	Electrical industrial equipment and apparatus
54	Household appliances
55	Electric lighting and wiring equipment
56	Audio, video, and communication equipment
57	Electronic components and accessories
58	Miscellaneous electrical machinery and supplies
59	Motor vehicles, truck and bus bodies, trailers, and motor vehicles parts
60	Aircraft and parts
61	Other transportation equipment
62	Scientific and controlling instruments
63	Ophthalmic and photographic equipment
64	Miscellaneous manufacturing
65	Railroads, motor freight, water and air transportation, pipelines
66	Communications, except radio and TV
67	Radio and TV broadcasting
68	Electric services, gas distribution, water and sanitary services
69	Retail and wholesale
70	Finance and insurance
71	Owner-occupied dwellings, real estate and royalties
72	Hotels and lodging places, personal and repair services except auto
73	Computer, legal, engineering, and accounting services, and advertising
74	Eating and drinking places
75	Automotive repair and services
76	Amusements
77	Health, educational and social services, and membership organizations

Table IA.VIII

BEA Industry - SIC Code Dictionary

BEA Industry	SIC Code
1+2	100-299
3	800-849, 860-919, 930-999
4	700-739, 750-799, 850-859, 920-929
5+6	1000-1079, 1090-1099
7	1200-1239, 1250-1299
8	1300-1379, 1390-1399
9+10	1400-1479, 1490-1499
11+12	1080-1089, 1240-1249, 1380-1389, 1480-1489, 1500-1799, 6550-6559
13	3480-3489, 3761, 3795
14	2000-2099, 5460-5469
15	2100-2199
16	2200-2249, 2260-2269, 2280-2289
17	2270-2279, 2290-2299
18	2250-2259, 2300-2389
19	2390-2399
20+21	2400-2499
22	2500-2519
23	2520-2599
24	2600-2649, 2660-2699
25	2650-2659
26	2700-2799
27	2800-2819, 2860-2899
28	2820-2829
29	2830-2849
30	2850-2859
31	2900-2999
32	3000-3099
33+34	3100-3199
35	3200-3229
36	3230-3299
37	3300-3329, 3390-3399, 3462
38	3330-3389, 3460-3461, 3463-3469
39	3400-3419
40	3430-3449
41	3450-3469
42	3420-3429, 3470-3479, 3490-3499
43	3500-3519
44	3520-3529
45	3530-3533
46	3534-3539
47	3540-3549
48	3550-3559
49	3560-3569
50	3590-3599
51	3570-3579

BEA Industry	SIC Code
52	3580-3589
53	3600-3629
54	3630-3639
55	3640-3649
56	3650-3669
57	3670-3679
58	3680-3699
59	3700-3715, 3717-3719
60	3720-3729, 3760, 3762-3769
61	3716, 3730-3759, 3770-3794, 3796-3799
62	3800-3849
63	3850-3899
64	3900-3999
65	4000-4299, 4400-4799
66	4800-4829, 4840-4899
67	4830-4839
68	4900-4999
69	5000-5459, 5470-5799, 5900-5999
70	6000-6499, 6700-6731, 6733-6799
71	6500-6549, 6560-6599
72	7000-7099, 7200-7299, 7600-7689
73	7300-7399, 7690-7699, 8100-8199, 8700-8732, 8734-8799
74	5800-5899
75	7500-7599
76	7800-7999
77	740-749, 6732, 8000-8099, 8200-8499, 8600-8699, 8733, 8800-8999