NYU STERN

Salomon Center for the Study of Financial Institutions

Working Paper Series **ASSET MANAGEMENT** Research Group

INSIDE INFORMATION AND THE OWN COMPANY STOCK PUZZLE

Stijn Van Nieuwerburgh
Laura Veldkamp

SC-AM-05-07

Inside Information and the Own Company Stock Puzzle

Stijn Van Nieuwerburgh and Laura Veldkamp *
New York University, Stern School of Business

September 20, 2005

Abstract

U.S. investors allocate 30-40% of their financial asset portfolio in the stock of the company stock they work for. Such a portfolio flies in the face of standard portfolio theory, which prescribes that an investor should hold less of a financial asset that is positively correlated with her undiversified labor income. Nevertheless, we propose a rational explanation that prescribes a long position in own company stock. Precisely because the own company stock is positively correlated with the investor's labor income, any information the investor learns about her earnings is a partial information advantage in her own company stock. When confronted with a choice of what information to acquire, employees may choose to learn about their own firm. Learning lowers the employee's risk of holding own-firm equity, which raises its risk-adjusted returns and makes a long position optimal.

*Stijn Van Nieuwerburgh: svnieuwe@stern.nyu.edu, NYU Stern, Finance Department, 44 West 4th St., 9th floor, New York, NY 10012. Laura Veldkamp: lveldkam@stern.nyu.edu, NYU Stern, Economics Department, 44 West 4th St., 7th floor, New York, NY 10012. JEL classification: F30, G11, D82. Keywords: Own company stock, asymmetric information, information theory.

Conventional wisdom dictates that an employee should take small or negative positions in her own company's equity, in order to hedge labor income risk (Baxter and Jermann (1997)). However, there is another way to reduce that risk: learn about future labor income realizations and adjust work effort accordingly. Learning is a substitute for hedging. Yet while hedging motives make own company stock less attractive, learning makes iy more attractive. Learning creates a private information advantage that induces an employee to hold more company stock, on average. We investigate how the incentives to hedge and to learn compete, and show that holding own-firm stock may not be so puzzling after all.

U.S. employees hold a large fraction of their portfolio in their own company's stock. Mitchell and Utkus (2002) document that 29% of assets in defined contribution plans are invested in own company stock.When direct and indirect ownership in own company stock is accounted for, 40% of directly and indirectly held public equity is invested in own company stock (Moskowitz and Vissing-Jorgensen (2001)). This behavior challenges standard portfolio theory, because an investor's human wealth is undiversified and tends to covary with payoffs to company stock; both depend on company performance. A diversified portfolio would therefore contain a small long or even short position in company stock.

The fact that labor income is a large risk makes both diversification and learning more valuable. Each has an opposite effect on her optimal portfolio. We model an investor who faces uncertainty about the payoffs from work as well as the payoffs from her financial portfolio. The latter consists of own company stock and other financial assets ('the market'). The employee has a fixed capacity to learn about future labor income and/or future income from the market asset, before choosing her work effort and asset portfolio. We characterize conditions under which she prefers to use all capacity to reduce uncertainty about labor income. Income information is particularly valuable because it enables the employee to adjust work effort and minimize a large risk in her total wealth portfolio. Because labor income and the payoff to company stock are positively correlated, the employee who learns about labor income reduces uncertainty about her company's stock in the process. This reduction in risk effectively increases her *risk-adjusted* own company stock return. Her optimal portfolio tilts towards a long position in own company stock. We show that inside information can overwhelm the competing desire to hold a portfolio that insures against labor income risk.

Modeling learning choices, rather than endowing employees with an information advantage, helps to explain cross-sectional variation in own company stock holdings. Employees of small and stand alone firms hold more company stock than employees of larger and conglomerate firms. If employees move from stand alone firms to conglomerates, they reduce own company stock (Cohen (2004)). Our model rationalizes these facts. If small firms' labor income covaries more with stock payoffs, then the ability to make better inference about company stock makes employees want to hold more of it.

We concur with Massa and Simonov (2005) who argue that loyalty (Cohen (2004)), familiarity (Huberman (2001), Hong, Kubik and Stein (2004)), and ambiguity (Boyle, Uppal and Wang (2003)) capture information advantages. They find that familiarity affects less-informed investors more, diminishes when the profession or location of the investor changes, and generates higher returns. By explicitly introducing an information choice, our theory can explain where these information advantages come from.

1 Model

This is a one-shot decision problem for an employee.[1] In period 1, the employee makes her learning choice. She allocates a fixed amount of precision between two signals: one about her wage, and one about the payoff of the market asset. In period 2, she observes her chosen signals and makes her investment choice. In period 3, she receives the asset payoffs and her wage, and consumes.

Preferences In order to study information acquisition, we want to start with investors who have a preference for early resolution of uncertainty. Investors, with absolute risk aversion parameter ρ, maximize their expected certainty equivalent wealth:[2]

$$U = E_1 \left\{ -\log \left(E_2[exp(-\rho W)] \right) \right\}. \tag{1}$$

The term $-\log \left(E_2[exp(-\rho W)] \right)$ is the level of consumption that makes the investor indifferent between consuming that amount for certain and investing in her optimal portfolio, in period 2. This certainty equivalent consumption is conditional on the realization of the employee's signals. Since these signals are not known in period 1, the investor maximizes the expected period-2 certainty equivalent, conditioning on information in prior beliefs.

Budget Constraint Let $r > 1$ be the risk-free return and (q_m, q_c) and (p_m, p_c) be the number of shares the investor chooses to hold and the asset prices of market assets (m) and company assets (c). Investor's terminal wealth is then her initial wealth W_0, plus the

[1]See VanNieuwerburgh and Veldkamp (2005b) for how to embed this in a general equilibrium model with a continuum of atomless investors.

[2]Utility can instead be defined over consumption by assuming that all wealth is consumed at the end of period 3. This formulation of utility has the desirable feature that it treats learned information and prior information as equivalent. It does so without losing the exponential structure of preferences that will keep the problem tractable.

profit she earns on her portfolio investments, plus the fixed component of labor income $\bar{\omega}$, and its variable component which takes the form of a bonus f_ω (net of the cost of effort), which the agent will receive if she exerts effort (ℓ):

$$W = rW_0 + q_m(f_m - p_m r) + q_c(f_c - p_c r) + \bar{\omega} + \ell f_\omega \tag{2}$$

To keep the model simple, effort is a binary choice $\ell \in \{0, 1\}$.

Initial information The employee is endowed with normally distributed prior beliefs about the payoff of the market $f_m \sim N(\mu_m, \sigma_m^2)$, the payoff of own company stock $f_c \sim N(\mu_c, \sigma_c^2)$ and the amount of her bonus $f_\omega \sim N(\mu_\omega, \sigma_\omega^2)$. For simplicity, the market payoff f_m is independent of the bonus f_ω and the own company stock payoff f_c.[3] This stacks the deck against us: making market payoffs independent of labor income creates strong incentives to diversify.

To capture the idea that holding company stock is a bad way to diversify labor income risk, f_ω and f_c are correlated. Both load on a common factor γ:

$$f_\omega = \mu_\omega + \beta\gamma + \epsilon \quad \text{and} \quad f_c = \mu_c + \gamma. \tag{3}$$

The bonus contains an idiosyncratic component $\epsilon \sim N(0, \sigma_\varepsilon^2)$, orthogonal to $\gamma \sim N(0, \sigma_\gamma^2)$. This structure imposes restrictions on the relationships between variances in the model: $\sigma_\omega^2 = \beta^2\sigma_\gamma^2 + \sigma_\epsilon^2$, $\sigma_c^2 = \sigma_\gamma^2$, and $cov(f_c, f_\omega) \equiv \xi = \beta\sigma_\gamma^2$.

[3] VanNieuwerburgh and Veldkamp (2005a) show how to set up a related problem when payoffs are correlated.

Information Acquisition At time 1, the employee chooses how much to learn about the market payoff and her bonus. She chooses the precision of two signals: $\eta_m \sim N(f_m, \sigma^2_{\eta m})$ and $\eta_\omega \sim N(f_\omega, \sigma^2_{\eta\omega})$. Because of Bayes' law, we can bypass the choice of signals, and model the choice over the posterior beliefs directly. An investor's posterior belief about the payoffs f_i, $i = m, \omega$, conditional on a prior belief μ_i and signal η_i, is formed from:

$$\hat{\mu}_i \equiv E[f_i|\mu_i, \eta_i] = \left(\sigma_i^{-2} + \sigma_{\eta i}^{-2}\right)^{-1}\left(\sigma_i^{-2}\mu_i + \sigma_{\eta i}^{-2}\eta_i\right) \tag{4}$$

with variance that is a harmonic mean of the signal variances:

$$\hat{\sigma}_i^2 \equiv V[f_i|\mu_i, \eta_i] = \left(\sigma_i^{-2} + \sigma_{\eta i}^{-2}\right)^{-1}. \tag{5}$$

In the appendix we also derive updating formulas for $\hat{\sigma}^2_\varepsilon$ and $\hat{\sigma}^2_\gamma$ using the Kalman filter.

There are 2 constraints governing how the investor can choose her signals about risk factors. The first is the *capacity constraint.* Capacity K can then be interpreted as the percentage by which an investor can decrease the risk she faces, where risk is measured as the generalized standard deviation of asset payoffs and labor income: $\hat{\sigma}^2_m \hat{\sigma}^2_\omega \geq e^{-2K}\sigma^2_m \sigma^2_\omega$.

This capacity constraint is one possible description of a learning technology. We think it is a relevant constraint because it is a commonly-used distance measure in econometrics (a log likelihood ratio); it has a long history as a quantity measure in information theory (Shannon (1948)); it is a measure of information complexity (Cover and Thomas (1991)), and it has been used to describe limited information processing ability in economic settings (Sims (2003)).

The second constraint is the *no negative learning constraint*: the investor cannot acquire signals that transmit negative information. We rule this out by requiring the variance of both signals to be positive. This implies that the posterior beliefs have a variance that is not greater than the prior beliefs: $\hat{\sigma}_m^2 \leq \sigma_m^2$ *and* $\hat{\sigma}_\omega^2 \leq \sigma_\omega^2$.

2 Results: Learning and Portfolio Choices

To solve the model, we work backwards. At time 2 the investor chooses her optimal asset portfolio (q_m, q_c) and her optimal work effort $\ell \in \{0, 1\}$, taking as given the posterior means $\hat{\mu}_i$ and variances $\hat{\sigma}_i$, where $i = \{m, c, \omega\}$.

Optimal Portfolio Choice Substituting the budget constraint (2) into the objective function (1), dropping the constant multiplier $(\rho r W_0 + \rho\bar{\omega})$, and taking period-2 expectations of a log normal variable delivers:

$$E_1\left[\rho\left\{q_m(\hat{\mu}_m - p_m r) + q_c(\hat{\mu}_c - p_c r) + \ell\hat{\mu}_\omega\right\} - \frac{\rho^2}{2}\left\{q_m^2\hat{\sigma}_m^2 + q_c^2\hat{\sigma}_c^2 + \ell^2\hat{\sigma}_\omega^2 + 2q_c\ell\hat{\xi}\right\}\right]$$

First order conditions with respect to q_m and q_c of the terms inside the expectation delivers the optimal portfolio rules:

$$q_m = \frac{1}{\rho}\hat{\sigma}_m^{-2}(\hat{\mu}_m - p_m r) = \frac{1}{\sigma_m\rho}\left(\frac{\hat{\mu}_m - p_m r}{\sigma_m}\right)y_m \quad (6)$$

$$q_c = \frac{1}{\rho}\hat{\sigma}_c^{-2}(\hat{\mu}_c - p_c r) - \hat{\sigma}_c^{-2}\ell\hat{\xi} = \frac{1}{\sigma_c\rho}\left(\frac{\hat{\mu}_c - p_c r}{\sigma_c}\right)y_c - \ell\beta\left(\frac{y_c}{y_\omega}\right) \quad (7)$$

where $y_i = \hat{\sigma}_i^{-2}/\sigma_i^{-2}$ is the proportional increase in belief precision for $i = m, c, \omega$. The appendix shows that the covariance between company payoffs and the bonus, conditional on the signal η_ω, is $\hat{\xi} = \beta\sigma_\gamma^2 y_\omega^{-1}$. Learning about wages $(y_\omega > 1)$ lowers its conditional covariance with company stock payoffs. This is one reason that the optimal holdings of company stock rise.

Optimal Work Effort Choice The employee exerts high effort ($\ell = 1$) iff[4]

$$\hat{\mu}_\omega - \beta(\hat{\mu}_c - p_c r)\left(\frac{y_c}{y_\omega}\right) - \frac{\rho}{2}\left(\hat{\sigma}_\omega^2 - \hat{\sigma}_\gamma^2 \beta^2 \left(\frac{y_c}{y_\omega}\right)^2\right) > 0. \tag{8}$$

The term involving posterior beliefs $\hat{\mu}_\omega - \beta\hat{\mu}_c$ is the only component of the labor decision not known at time 1. It is a function of constants and the posterior beliefs $\beta\hat{\gamma}\,(1 - y_c/y_\omega) + \hat{\varepsilon}$, which are determined by the value of the observed signal η_ω. Thus, the employee exerts high effort when her signal about the bonus is above a cutoff. Condition (8) is equivalent to $D(y_\omega)\eta_\omega > C(y_\omega)$, where $D(y_\omega) = \left(1 - (y_c/y_\omega)(\beta^2\sigma_\gamma^2/\sigma_\omega^2)\right)(1 - y_\omega^{-1})$, and the cutoff $C(y_\omega) = -\mu_\omega + \sigma_\omega^2(\beta(\mu_c - p_c r) + \sigma_\varepsilon^2 \rho/2)/(\sigma_\varepsilon^2 y_\omega + \beta^2\sigma_\gamma^2)$ (see appendix for details).

Assumption 1. *The agent is indifferent between working and not working, given prior information only:* $C(1) = 0$*, or equivalently,* $\mu_\omega = \beta(\mu_c - p_c r) + \frac{\rho}{2}\sigma_\varepsilon^2$*.*

The assumption equates the prior expected bonus μ_ω with the cost of labor income risk. When the employee hedges γ-risk by holding less own-company stock, $\beta(\mu_c - p_c r)$ is the asset payoff she forgoes. The idiosyncratic ε-risk cannot be hedged, and has utility cost $\frac{\rho}{2}\sigma_\varepsilon^2$. Assumption 1 makes the agent ex-ante uncertain about whether exerting high effort is optimal. This avoids a situation where information about labor income is not valuable because the employee is (almost) certain that she will or will not work.

Optimal Information Choice At time 1, when information is chosen, signals $\{\eta_m, \eta_\omega\}$, and thus labor and portfolio choices are random variables. We define three pieces of new

[4] This result follows from substituting the optimal portfolios back into the utility function: $E_1\left[\frac{1}{2}\left(\frac{\hat{\mu}_m - p_m r}{\sigma_m}\right)^2 y_m + \frac{1}{2}\left(\frac{\hat{\mu}_c - p_c r}{\sigma_c}\right)^2 y_c + \rho\ell\left(\hat{\mu}_\omega - (\hat{\mu}_c - p_c r)\beta\left(\frac{y_c}{y_\omega}\right)\right) - \frac{\rho^2\ell^2}{2}\left(\hat{\sigma}_\omega^2 - \hat{\sigma}_\gamma^2\beta^2\left(\frac{y_c}{y_\omega}\right)^2\right)\right]$.

notation. Let prior Sharpe ratios be $\theta_i = \frac{\mu_i - p_i r}{\sigma_i}$ for $i = m, c$. Let the standard deviation of $D(y_\omega)\eta_\omega$ (a mean-zero normal variable) be denoted $\sigma_D = D(y_\omega)(\sigma_\omega^2 + \sigma_{\eta\omega}^2)^{.5}$. Let the probability that the high effort is chosen be denoted $\Pi = 1 - \Phi(C(y_\omega)/\sigma_D)$, where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the standard normal c.d.f. and p.d.f.. The time-1 problem then is to choose $\{y_m, y_\omega\}$ to maximize the expected utility:

$$EU(y_\omega, y_m) = \frac{1}{2} y_m (1 + \theta_m^2) + \frac{1}{2} y_c (1 + \theta_c^2) + \rho \left(\phi \left(\frac{C(y_\omega)}{\sigma_D} \right) \sigma_D - C(y_\omega) \right) \Pi.$$

subject to the capacity constraint $y_m y_\omega \leq e^{2K}$ and the no-forgetting constraints $y_m \geq 1$, $y_\omega \geq 1$. The first two terms represent expected portfolio holdings, times payoffs, which are squared Sharpe ratios, with mean equal to $1 + \theta_i^2$. The third term is the posterior expected value of working hard $E_1 \left[\ell \left(D(y_\omega)\eta_\omega - C(y_\omega) \right) \right]$. We can think of there being one choice variable y_ω, the capacity devoted to learning about one's bonus f_ω, with $y_m = e^{2K} y_\omega^{-1}$ because the capacity constraint holds with equality.

The marginal value of information about the bonus consists of five terms:

$$MU_\omega(y_\omega; \mu_\omega, \mu_c, p_c, r, \beta, \rho, \sigma_\varepsilon, \sigma_\gamma) =$$

$$\frac{1}{2}(1 + \theta_m^2)\frac{\partial y_m}{\partial y_\omega} + \frac{1}{2}(1 + \theta_c^2)\frac{\partial y_c}{\partial y_\omega} + \rho \frac{\partial \Pi \sigma_D \phi \left(\frac{C(y_\omega)}{\sigma_D} \right)}{\partial y_\omega} - \rho C(y_\omega) \frac{\partial \Pi}{\partial y_\omega} - \rho \Pi \frac{\partial C(y_\omega)}{\partial y_\omega}$$

The first term represents the foregone benefit of using capacity to reduce uncertainty about the market asset. The second term shows that the higher the squared Sharpe ratio of the agent's own company stock, the more profit she can make by using her inside information to take large long or short positions in the stock. The third term captures the benefit of working hard er when bonuses are larger. The last two terms measure the positive effect of learning on the probability of working hard. When the agent learns more about her bonus, working hard becomes less risky, and thus more desirable. (See appendix for proof.)

The next two propositions show when it is optimal to learn about the bonus instead of the market. This depends on whether the objective function is convex or concave. The appendix gives the conditions for convexity and concavity. Numerical examples described below show that the objective function is typically convex.

Proposition 1. *If the objective is convex in y_ω, and if $EU(e^{2K}, 1) > EU(1, e^{2K})$, then the optimal strategy is to use all capacity to learn about f_ω.*

Proposition 2. *If the objective function is concave and (i) if $MU(1; \cdot) > 0$, then the optimal information acquisition strategy uses some capacity to learn about f_ω; (ii) if $MU(e^{2K}; \cdot) > 0$, then the optimal information acquisition strategy is to use all capacity to learn about f_ω.*

Own Company Stock Bias The optimal portfolio (6) and (7) depends on the random signal realization in period two. We therefore characterize own company bias in the expected portfolios.

Proposition 3. *The optimal expected asset portfolio is $E_1[q_m] = y_m\theta_m/(\rho\sigma_m)$ and $E_1[q_c] = y_c\theta_c/(\rho\sigma_c) - \beta\Pi y_c/y_\omega$.*

For comparison, consider the portfolio allocation in the benchmark model without learning capacity ($K = 0$): $E_1[q_m^{no\,learn}] = \theta_m/(\rho\sigma_m)$ and $E_1[q_c^{no\,learn}] = \theta_c/(\rho\sigma_c) - \beta\Pi^{no\,learn}$. When labor income payoffs are positively correlated with own company stock payoffs ($\beta > 0$), the benchmark model prescribes a lower position in the own company stock ($-\beta\Pi^{no\,learn} < 0$). This is the hedging effect of Baxter and Jermann (1997).

When employees can learn ($K > 0$), and they choose to learn about labor income ($y_\omega > 1$), two effects operate on expected portfolios. The first is an information effect:

learning about labor income provides information about company stock. This leads the employee to tilt her portfolio towards a *longer* position in own company stock (provided $\theta_c > 0$). The largest tilt occurs when all capacity is devoted to learning about labor income: $y_\omega = e^{2K}$. Second, learning affects the need to hedge labor income risk. On the one hand, learning causes the employee to work harder and generates more labor income to hedge, on average. This decreases $E[q_c]$. On the other hand, hedging is not as useful because some of that risk is already being resolved through learning. This increases $E[q_c]$. The net effect on the hedging component of the portfolio is ambiguous. We now illustrate these effects with a numerical example.

Numerical Example We set the prior mean-to-variance ratio of both assets equal so that in the benchmark case with riskless labor income and no signals, the investor would then hold an equal amount of each. We set $\rho = 3$, $\sigma_m = .20$, $\beta = .8$, $\sigma_\gamma = \sigma_c = .30$, $\sigma_\varepsilon = .05$. For these parameters, the Sharpe ratio on the market $\theta_m = .4$ and the Sharpe ratio on own-company stock is $\theta_c = .6$. We then give the employee enough capacity to eliminate 40% of the volatility in the market asset or in the bonus through learning ($K = .4$). Figure 1), left panel, shows that it is optimal to specialize in learning about the bonus (illustrates proposition 1). The right panel compares the expected portfolios across y_ω (proposition 3). Since the optimal choice is $y_\omega = e^{2K}$, we read off the optimal portfolios on the right of the graph. Relative to the benchmark no-learning case ($y_\omega = 1$), the employee optimally holds a long position in own-company stock, which is much higher than in the no-learning economy (57% versus 28%).

Cross-Sectional Patterns Cohen (2004) documents that employees of conglomerates allocate a smaller fraction of their discretionary 401(k) contributions to own company stock

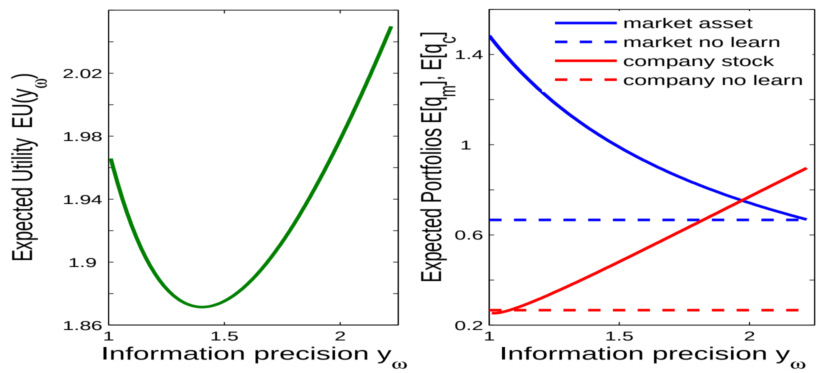

Figure 1: Expected Utility and Portfolio for Various Learning Choices ($\beta = 0.8$). The right panel plots expected holdings of market and own-firm assets held in our model (solid lines) and a no-learning economy (dashed lines). Since the utility maximizing learning choice is $y_\omega = e^{2K}$, the optimal learning portfolio is the amount of each asset at the intersection with the right axis. The parameters are as in the text. The employee has learning capacity of $K = 0.4$, so that the upper bound for y_ω is $e^{2K} = 2.23$.

than employees of stand-alone firms. He argues this is due to stronger loyalty to stand-alone firms. Our information-based story rationalizes this fact.

We model conglomerate firms as low-β firms: their employees' bonus is less correlated with company stock payoffs. A lower β affects portfolios in three ways. First, a lower β weakens the information effect: information about the bonus generates less information about own company stock ($\partial y_c/\partial\beta > 0$). More uncertainty makes holding own-company stock less desirable. Second, a lower β has an ambiguous effect on hedging ($-\beta\Pi y_c/y_\omega$). The lower conditional covariance of stock payoffs and labor income ($\partial\beta y_c y_\omega^{-1}/\partial\beta > 0$) makes the employee want to hedge less. But, when labor income becomes less risky (σ_ω^2 falls because $\beta^2\sigma_\gamma^2$ falls), the employee work harder ($\partial\Pi/\partial\beta < 0$) and has more expected income to hedge. Third, when β declines, the value of information about the bonus may decline enough so that the optimal learning strategy switches from specialization in learning about the bonus ($y_\omega = e^{2K}$) to specialization in learning about the market ($y_\omega = 1$). Figure 2 illustrates this. When $\beta = .4$ instead of .8, it is optimal to learn about the market ($y_\omega = 1$). Such a switch discretely reduces own company stock holdings to the same level as in the no-learning model.

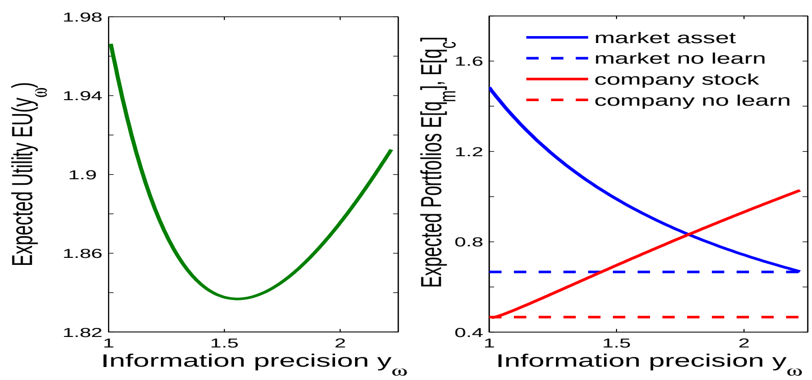

Figure 2: Expected Utility and Portfolios for Various Learning Choices ($\beta = 0.4$). Since the utility maximizing learning choice is $y_\omega = 1$, the optimal learning portfolio is the amount of each asset at the intersection with the left axis.

3 Conclusion

This paper explores learning about risky labor income as a rational explanation for the own company stock puzzle. It augments VanNieuwerburgh and Veldkamp (2005a) with labor income risk that can be resolved by hedging (holding less company stock), or by learning. Learning induces an employee to tilt her portfolio towards own company stock, while reducing the need to hedge labor income risk. Under conditions described in the paper, the opportunity to learn makes an own company stock 'bias' optimal.

References

Baxter, Marianne and Urban Jermann, "The International Diversification Puzzle Is Worse Than You Think," *The American Economic Review*, 1997, *87(1)*, 179–180.

Boyle, Phelim, Raman Uppal, and Tan Wang, "Ambiguity Aversion and the Puzzle of Own-Company Stock in Pension Plans," 2003. Working Paper.

Cohen, Lauren, "Loyalty Based Portfolio Choice," 2004. Working Paper.

Cover, Thomas and Joy Thomas, *Elements of Information Theory*, first ed., John Wiley and Sons, New York, New York, 1991.

Hong, Harrison, Jeffrey Kubik, and Jeremy Stein, "The Only Game in Town: Stock-Price Consequences of Local Bias," 2004. Working Paper.

Huberman, Gur, "Familiarity Breeds Investment," *Review of Financial Studies*, 2001, *14(3)*, 659–80.

Massa, Massimon and Andrei Simonov, "Hedging, Familiarity and Portfolio Choice," *The Review of Financial Studies*, 2005, *Forthcoming.*

Mitchell, Olivia and Stephen P. Utkus, "Company Stock and Retirement Plan Diversification," 2002. Pension Research Council Working Paper.

Moskowitz, Tobias and Annette Vissing-Jorgensen, "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?," *American Economic Review*, 2001, *92(4)*, 745–78.

Shannon, C.E., "Mathematical Theory of Communication," *Bell System Technology Journal*, 1948, *27*, 379–423 and 623–656.

Sims, Christopher, "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, *50(3)*, 665–90.

VanNieuwerburgh, Stijn and Laura Veldkamp, "Information Acquisition and Portfolio Under-Diversification," 2005. Working Paper.

___ **and** ___ , "Information Immobility and the Home Bias Puzzle," 2005. Working Paper.

A Technical Details

Posterior variances and covariances We start by deriving the relationship between precision ratios $y_c \equiv \hat{\sigma}_\gamma^{-2}/\sigma_\gamma^{-2}$ and $y_\omega \equiv \hat{\sigma}_\omega^{-2}/\sigma_\omega^{-2}$. The posterior precision about the bonus is the prior precision plus the signal precision: $y_\omega = 1 + \sigma_{\eta\omega}^{-2}/\sigma_\omega^{-2}$. Therefore $\sigma_{\eta\omega}^2 = (y_\omega - 1)^{-1}\sigma_\omega^2$. The posterior variances $\hat{\sigma}_\epsilon^2$ and $\hat{\sigma}_\gamma^2$ are obtained from the Kalman filter:

$$\hat{\sigma}_\epsilon^2 = \sigma_\epsilon^2\left(\frac{\beta^2\sigma_\gamma^2 + \sigma_{\eta\omega}^2}{\sigma_\epsilon^2 + \beta^2\sigma_\gamma^2 + \sigma_{\eta\omega}^2}\right) = \sigma_\epsilon^2\left(1 - \frac{\sigma_\epsilon^2}{\sigma_\omega^2}(1 - y_\omega^{-1})\right), \tag{9}$$

$$\hat{\sigma}_\gamma^2 = \sigma_\gamma^2\left(\frac{\sigma_\epsilon^2 + \sigma_{\eta\omega}^2}{\sigma_\epsilon^2 + \beta^2\sigma_\gamma^2 + \sigma_{\eta\omega}^2}\right) = \sigma_\gamma^2\left(1 - \frac{\beta^2\sigma_\gamma^2}{\sigma_\omega^2}(1 - y_\omega^{-1})\right) \tag{10}$$

Rearranging (10), the extra precision about the own company stock payoff, inferred from the signal about the bonus is: $y_c = (\sigma_\epsilon^2 + \beta^2\sigma_\gamma^2)/(\sigma_\epsilon^2 + \beta^2\sigma_\gamma^2 y_\omega^{-1})$. Note that if $y_\omega = 1$, then $y_c = 1$. Learning more about the bonus ($y_\omega > 1$) increases the posterior precision of the own company stock ($y_c > 1$), but not by the same amount:

$$\frac{\partial y_c}{\partial y_\omega} = \beta^2\sigma_\gamma^2 \frac{\sigma_\epsilon^2 + \beta^2\sigma_\gamma^2}{[\sigma_\epsilon^2 y_\omega + \beta^2\sigma_\gamma^2]^2} \in [0,1],$$

Learning more about future labor income increases the own company asset precision at a decreasing rate ($\partial^2 y_c/\partial y_\omega^2 < 0$). Therefore, there are decreasing returns in the expected investment profit to learning about labor information.

Conditional on seeing the signal η_ω, the (posterior) covariance between f_c and f_ω is $\hat{\xi} = \beta\hat{\sigma}_\gamma^2 + Cov(\gamma, \varepsilon|\eta_\omega)$. Combining y_ω, (9), and (10) and the formula for the variance of a sum, yields $\hat{\xi} = \beta\sigma_\gamma^2 y_\omega^{-1}$.

Work Effort Choice Using expressions in (3), condition (8) can be written as $\beta\hat{\gamma}(1 - y_c/y_\omega) + \hat{\varepsilon} > C(y_\omega)$, where

$$C(y_\omega) = \frac{\rho}{2}\left(\hat{\sigma}_\omega^2 - \beta^2\hat{\sigma}_\gamma^2\left(\frac{y_c}{y_\omega}\right)^2\right) - \mu_\omega + \beta(\mu_c - p_c r)\left(\frac{y_c}{y_\omega}\right).$$

$C(y_\omega)$ can then be manipulated to yield the expression in the text. We posit $\beta\hat{\gamma}(1 - y_c/y_\omega) + \hat{\varepsilon} = D(y_\omega)\eta_\omega$ and use the Kalman filtering formulas $\hat{\varepsilon} = (\sigma_\varepsilon^2\eta_\omega)/(\sigma_\omega^2 + \sigma_{\eta\omega}^2)$ and $\hat{\gamma} = (\beta^2\sigma_\gamma^2\eta_\omega)_{/(\sigma_\omega^2+\sigma_{\eta\omega}^2)}$ to solve for $D(y_\omega)$. The standard deviation of $D(y_\omega)\eta$ is $\sigma_D = D(y_\omega)(\sigma_\omega^2 + \sigma_{\eta\omega}^2)^{.5} = (\sigma_\omega\sigma_\varepsilon^2(y_\omega^2 - y_\omega)^{.5})/(\sigma_\varepsilon^2 y_\omega + \beta^2\sigma_\gamma^2)$.

Evaluating derivatives Two key derivatives needed to evaluate the marginal utility w.r.t. y_ω are:

$$\frac{\partial C(y_\omega)}{\partial y_\omega} = -\frac{\sigma_\omega^2\sigma_\varepsilon^2}{(\sigma_\varepsilon^2 y_\omega + \beta^2\sigma_\gamma^2)^2}\left(\beta(\mu_c - p_c r) + \frac{\rho}{2}\sigma_\varepsilon^2\right) < 0$$

$$\frac{\partial \sigma_D}{\partial y_\omega} = \frac{\sigma_\omega\sigma_\varepsilon^2(.5\sigma_\varepsilon^2 y_\omega + \beta^2\sigma_\gamma^2(y_\omega - .5))}{(y_\omega^2 - y_\omega)^{.5}(\sigma_\varepsilon^2 y_\omega + \beta^2\sigma_\gamma^2)^2} > 0$$

Define $x \equiv C(y_\omega)/\sigma_D < 0$, negative by assumption 1. Combining all terms of the marginal utility of y_ω:

$$MU_\omega(y_\omega; \mu_\omega, \mu_c, p_c, r, \beta, \rho, \sigma_\varepsilon, \sigma_\gamma) = \frac{1}{2}(1 + \theta_c^2)\frac{\partial y_c}{\partial y_\omega} + \rho\phi(x)\frac{\partial \sigma_D}{\partial y_\omega} - \rho\Pi\frac{\partial C(y_\omega)}{\partial y_\omega} > 0 \tag{11}$$

where

$$\frac{\partial x}{\partial y_\omega} = \frac{1}{\sigma_D}\left[\frac{\partial C(y_\omega)}{\partial y_\omega} - x\frac{\partial \sigma_D}{\partial y_\omega}\right]$$

$$\frac{\partial \Pi}{\partial y_\omega} = -\phi(x)\frac{\partial x}{\partial y_\omega}$$

$$\frac{\partial \Pi E[D(y_\omega)\eta_\omega | D(y_\omega)\eta_\omega \geq C(y_\omega)]}{\partial y_\omega} = \phi(x)\left[\left(1 + x^2\right)\frac{\partial \sigma_D}{\partial y_\omega} - x\frac{\partial C(y_\omega)}{\partial y_\omega}\right].$$

Since $\partial C(y_\omega)/\partial y_\omega < 0$, $\mathrm{MU}(y_\omega)$ is always positive. The total marginal utility w.r.t. y_ω also contains the term $-\frac{1}{2}(1+\theta_m^2)e^{2K}y_\omega^{-2}$, which measures how y_m changes as the employee increases y_ω, through the capacity constraint $y_m y_\omega = e^{2K}$.

Concavity, convexity and optimal y_ω The objective function is convex when the following second derivative of utility w.r.t y_ω is positive, and concave when it is negative. The expression is a long, but straightforward function of y_ω and parameters only.

$$\frac{1}{2}(1+\theta_c^2)\frac{\partial^2 y_m}{\partial y_\omega^2} + \frac{1}{2}(1+\theta_c^2)\frac{\partial^2 y_c}{\partial y_\omega^2} + \rho\phi(x)\frac{\partial^2 \sigma_D}{\partial y_\omega^2} - \rho\Pi\frac{\partial^2 C(y_\omega)}{\partial y_\omega^2} + \rho\phi(x)\frac{1}{\sigma_D}\left[\frac{\partial C(y_\omega)}{\partial y_\omega} - x\frac{\partial \sigma_D}{\partial y_\omega}\right]^2.$$

The first term is positive: $\partial^2 y_c/\partial y_\omega^2 = 2e^{2K}y_\omega^{-3}$. We already showed that y_c is concave in y_ω: $\partial^2 y_c/\partial y_\omega^2 < 0$ and the second term is negative. It can be shown that $\partial^2\sigma_D/\partial y_\omega^2 < 0$ and $\partial^2 C(y_\omega)/\partial y_\omega^2 > 0$, so that the third and fourth terms are negative as well. The last term is clearly positive.

If the objective function is convex $\forall y_\omega \in (1, e^{2K})$, full specialization always arises. It is optimal to learn about the bonus if the objective is higher at $y_\omega = e^{2K}$ then at $y_m = e^{2K}$ (proposition 1). If the objective is concave, and $MU_\omega(1;\cdot) > 0$, then it is optimal to allocate the first increment of capacity towards learning about the bonus instead of the market (proposition 2, first part). If also $MU_\omega(e^{2K};\cdot) > 0$, then allocating the last increment of capacity to learning about the bonus is still more valuable than allocating it to learning about the market, and full specialization in learning about the bonus takes place (proposition, second part).