Internet Appendix: Learning about the Consumption Risk Exposure of Firms

Yongjin Kim
Anderson School of Management
University of New Mexico

Lars-Alexander Kuehn
Tepper School of Business
Carnegie Mellon University

Kai Li
HSBC Business School
Peking University

October 26, 2023

In the Internet Appendix, we examine the robustness of our empirical results when idiosyncratic productivity shocks are heteroskedastic (Appendix A), when we use alternative measurements of investment rates and Q in case of missing PPENT (Appendix B), and when a longer burn-in period is imposed in the learning (Appendix C). In addition, we compare Bayesian learning with a frequentist approach (Appendix D) and learning with fading memory (Appendix E).

A Heteroskedasticity in Idiosyncratic Shocks

In this section, we present the learning about risk exposure when idiosyncratic shocks to productivity are heteroskedastic. With time-varying idiosyncratic risk, the productivity growth process becomes

$$g_{i,t+1} = \mu + b(\sigma_c \eta_{t+1}) + \sigma_t \varepsilon_{i,t+1}, \tag{A.1}$$

where σ_t is the conditional volatility of idiosyncratic shock. With the revised dynamics, the beliefs about risk exposure b are updated according to

$$m_{b,t} = (1 - \kappa_t \sigma_{b,t}^2) m_{b,t-1} + \kappa_t \sigma_{b,t}^2 \hat{b}_t \tag{A.2}$$

$$\frac{1}{\sigma_{b,t}^2} = \frac{1}{\sigma_{b,t-1}^2} + \kappa_t, \tag{A.3}$$

where $\kappa_t = n_t \eta_t^2 \sigma_c^2 / \sigma_t^2 \geq 0$ and $\hat{b}_t = \left[\sigma_c \eta_t \sum_{i=1}^{n} (g_{i,t} - \mu)\right] / (n_t \sigma_c^2 \eta_t^2)$. Importantly, the evolution of mean beliefs (A.2) depends on the informativeness of the data κ_t. When idiosyncratic risk σ_t increases, the data becomes less informative, and the sample estimate $\hat{b}_t$ receives a lower weight in the updating.

To account for the impact of time-varying idiosyncratic volatility on parameter learning, we estimate the conditional volatility of idiosyncratic shocks for each industry as follows.

Step 1: We run a pooled panel regression of firm-level productivity onto consumption and obtain the residuals.

Step 2: For each year, we calculate the standard deviation of the cross-sectional residuals. Repeating this calculation year by year, we obtain the time series of the standard deviation.

Step 3: We fit the time series using a smoothing spline and use the fitted value each year as the estimate of the conditional volatility σ_t.

Having estimated the time series of the conditional volatility for each industry, we use σ_t in the learning equations (A.2) and (A.3) and obtain an alternative estimate of the risk exposure $m_{b,t}^{\text{dyn vol}}$. Next, we test whether these beliefs affect investment, Q, and the cost of capital.

We find that firm variables continue to respond strongly to the revised Bayesian beliefs in Table A1. Specifically, $m_{b,t}^{\text{dyn vol}}$ negatively predicts the investment rate and Q at the 1% significance level. Also, the risk estimate relates positively with the implied cost of capital, although the link is statistically weak. Overall, the alternative estimate's predictive power is largely comparable to our baseline estimates, suggesting that our empirical findings are robust to heteroskedasticity.

B Extended Measurements of Investment Rates and Q

Our measurement of investment rates and Q depends on the availability of the variable PPENT in the Compustat dataset; however, this value is missing for 23% firm-year observations. By imposing the filter on PPENT, we cover 77% of the pooled firm-year observations or 86% of the market capitalization of the full sample regarding the investment regressions.

In this section, we extend the measurement of investment rates and Q when PPENT is missing. Specifically, we impute missing investment rates by capital expenditure, less sale of PPE, plus intangible investment over total assets, i.e., (CAPX - SPPE + intangible investments)/AT. Similarly, we impute missing Q as the market equity plus book debt over

Table A1: Heteroskedasticity of Idiosyncratic Shocks

This table presents regressions of firm variables on risk exposure beliefs. Across six specifications, we regress investment rates $I_{i,t}/K_{i,t}$, $Q_{i,t}$, or the implied cost of capital $\text{ICC}_{i,t}$ on risk-exposure beliefs. $m_{b,t}$ is our baseline estimate with constant volatility of idiosyncratic shocks. $m_{b,t}^{\text{dyn vol}}$ is the alternative estimate with dynamic volatility of idiosyncratic shocks. Specifications (1) through (4) present panel regression results with industry fixed effects. Controls are size, age, profitability, and leverage; standard errors are clustered by firms. Specifications (5) and (6) present the Fama-MacBeth regression results. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Dep. variable	$I_{i,t}/K_{i,t}$		$Q_{i,t}$		$\text{ICC}_{i,t}$	
Specification	(1)	(2)	(3)	(4)	(5)	(6)
$m_{b,t}$	-0.0035***		-0.0176***		0.0035**	
	(-7.36)		(-2.75)		(2.14)	
$m_{b,t}^{\text{dyn vol}}$		-0.002***		-0.0225***		0.0004
		(-4.10)		(-3.03)		(0.38)
Controls	Yes	Yes	Yes	Yes	No	No
Industry FE	Yes	Yes	Yes	Yes	n.a.	n.a.
Year FE	No	No	No	No	n.a.	n.a.
N	112,155	112,155	121,410	121,410	91,486	91,486
adj. R^2	0.117	0.116	0.130	0.130	0.00005	-0.00002

total assets. Table A2 presents the regression results under this revision. We find that both investment rates and Q still respond negatively to risk exposure beliefs in the extended dataset.

C Bayesian Learning with Different Burn-In Periods

In our estimation of risk exposure beliefs, we assume five years of burn-in period to offset the influence of prior beliefs. To check the robustness of our findings, we increase the burn-in to 10 years and examine the empirical link between the revised risk estimates and firm variables. Table A3 shows that all of our findings are still highly significant.

D Frequentist Estimation

In our baseline estimation, we employ Bayesian learning to update the parameter beliefs with growing observations. Alternatively, in the spirit of a frequentist, we consider a linear regression model with an expanding window to utilize a growing dataset. Specifically, in the

Table A2: Robustness Check using Total Assets

This table presents regression results of investment rates $I_{i,t}/K_{i,t}$ and $Q_{i,t}$ on risk exposure beliefs $m_{b,t}$. We compute missing investment rates as capital expenditure, less sales of PPE, plus intangible investment over total assets, i.e., (CAPX - SPPE + intangible investment)/AT; and missing Q as market equity plus book debt over total assets. All specifications are panel regression with industry fixed effects. Controls are size, age, size, profitability, and leverage; standard errors are clustered by firms. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Dep. variable	$I_{i,t}/K_{i,t}$ (PPENT)	$I_{i,t}/K_{i,t}$ (AT)	$Q_{i,t}$ (PPENT)	$Q_{i,t}$ (AT)
Specification	(1)	(2)	(3)	(4)
$m_{b,t}$	-0.0035***		-0.0176***	
	(-7.36)		(-2.75)	
$m_{b,t}$		-0.0033***		-0.0443***
		(-6.76)		(-3.51)
Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	No	No	No	No
N	112,155	144,051	121,410	154,492
adj. R^2	0.117	0.137	0.130	0.101

productivity process

$$g_{i,t+1} = \mu + b(\sigma_c \eta_{t+1}) + \sigma \epsilon_{i,t+1}, \tag{D.1}$$

the risk exposure b can be identified through a linear regression of productivity $g_{i,t+1}$ onto the consumption shock $(\sigma_c \eta_{t+1})$. In this framework, we expand the estimation window every year to include new observations across time and firms, to update the alternative risk exposure $\hat{b}_t^{\text{expand}}$.

To examine the difference between Bayesian and frequentist estimates, we express the Bayesian estimate as follows:

$$m_{b,t} = (\sigma_{b,t})^2 \left[\frac{m_{b,0}}{\sigma_{b,0}^2} + \frac{1}{\sigma^2} \sum_{i=1}^{t} \bar{g}_i^\top (\bar{g}_{c,i} \mathbb{1}_{n_i}) \right] \tag{D.2}$$

$$= (\sigma_{b,t})^2 \left[\frac{m_{b,0}}{\sigma_{b,0}^2} + \frac{\sum_{i=1}^{t} n_i g_{c,i}^2}{\sigma^2} \underbrace{\left[\sum_{i=1}^{t} n_i g_{c,i}^2 \right]^{-1} \sum_{i=1}^{t} \bar{g}_i^\top (\bar{g}_{c,i} \mathbb{1}_{n_i})}_{=\hat{b}_t^{\text{expand}}} \right]. \tag{D.3}$$

This expression shows that the Bayesian estimate is in fact a weighted average of the prior

Table A3: Belief Estimates with Different Burn-in Periods

This table presents regressions of firm variables on risk exposure beliefs. Across six specifications, we regress investment rates $I_{i,t}/K_{i,t}$, $Q_{i,t}$, or the implied cost of capital $\text{ICC}_{i,t}$ on risk-exposure beliefs. $m_{b,t}$ is our baseline estimate with a burn-in of five years. As an alternative, we increase the burn-in to 10 years and obtain $m_{b,t}$ (10 year burn-in). Specifications (1) through (4) present panel regression results with industry fixed effects. Controls are size, age, profitability, and leverage; standard errors are clustered by firms. Specifications (5) and (6) present the Fama-MacBeth regression results. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Dep. variable	$I_{i,t}/K_{i,t}$		$Q_{i,t}$		$\text{ICC}_{i,t}$	
Specification	(1)	(2)	(3)	(4)	(5)	(6)
$m_{b,t}$ (5 years burn-in)	-0.0035***		-0.0176***		0.0035**	
	(-7.36)		(-2.75)		(2.14)	
$m_{b,t}$ (10 years burn-in)		-0.0031***		-0.0194***		0.0028*
		(-6.99)		(-3.29)		(1.76)
Controls	Yes	Yes	Yes	Yes	No	No
Industry FE	Yes	Yes	Yes	Yes	n.a.	n.a.
Year FE	No	No	No	No	n.a.	n.a.
N	112,155	111,562	121,410	120,987	91,486	87,495
adj. R^2	0.117	0.117	0.130	0.130	0.00005	0.00007

belief $m_{b,0}$ and the ordinary least square estimate $\hat{b}_t^{\text{expand}}$ from expanding regressions. The main distinction between the Bayesian and frequentist estimates is the presence of prior beliefs. In the Bayesian approach, a rational prior disciplines the estimate, especially in the early stage of learning when only a few observations are available. In contrast, the frequentist approach does not utilize prior beliefs. As a result, the early stage of estimation is driven by few observations, which could be very noisy. In fact, firm level productivity is very volatile, 92% annually according to our structural estimation, rendering the early sample useless. We counterbalance this effect by increasing the burn-in period to 15 years. In the limit, the Bayesian estimate converges to the frequentist one.

Table A4 shows the empirical association between the expanding window estimate and firm variables. As in the benchmark results, risk exposure estimates negatively predict investment rates and Q and positively the cost of capital. Due to the increase in noise in the expanding window regressions, the impact on the implied cost of capital is not significant at the 5% level. However, for investment rates and Q, it is.

Table A4: Beliefs from Expanding Window Regression

This table presents regressions of firm variables on risk exposure beliefs. Across six specifications, we regress investment rates $I_{i,t}/K_{i,t}$, $Q_{i,t}$, or the implied cost of capital $\text{ICC}_{i,t}$ on risk-exposure beliefs. $m_{b,t}$ is our baseline estimate. $\hat{b}_t^{\text{expand}}$ is the alternative belief that is obtained from regressions with expanding window. Specifications (1) through (4) present panel regression results with industry fixed effects. Controls are size, age, profitability, and leverage; standard errors are clustered by firms. Specifications (5) and (6) present the Fama-MacBeth regression results. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Dep. variable	$I_{i,t}/K_{i,t}$		$Q_{i,t}$		$\text{ICC}_{i,t}$	
Specification	(1)	(2)	(3)	(4)	(5)	(6)
$m_{b,t}$	-0.0035***		-0.0176***		0.0035**	
	(-7.36)		(-2.75)		(2.14)	
$\hat{b}_t^{\text{expand}}$		-0.0019***		-0.0189***		0.0043
		(-3.20)		(-5.21)		(1.27)
Controls	Yes	Yes	Yes	Yes	No	No
Industry FE	Yes	Yes	Yes	Yes	n.a.	n.a.
Year FE	No	No	No	No	n.a.	n.a.
N	112,155	93,563	121,410	103,391	91,486	87,495
adj. R^2	0.117	0.112	0.130	0.137	0.0005	-0.00002

E Learning with Fading Memory

A growing number of studies, including Malmendier and Nagel (2016) and Nagel and Xu (2022), document non-Bayesian learning in asset valuation that emerge from fading memory. Both Bayesian and non-Bayesian learning share the same insight that past observations shape the beliefs about parameters. In this section, we highlight the theoretical difference between the two learning forms.

The key difference lies in the aggregation of information. In Bayesian learning, every past observation has identical significance in forming parameter beliefs. In contrast, agents with fading memory put more weight on recent observations and thus the influence of past data on beliefs gradually fades away over time.

To elaborate on the comparison, let us consider the posterior probability of b in our model:

$$\text{Prob}(b|g_1,..,g_t,g_{c,1},..,g_{c,t}) \propto \text{Prob}(g_1,..,g_t|b,g_{c,1},..,g_{c,t}) \times \text{Prob}(b|g_{c,1},..,g_{c,t-1}), \quad \text{(E.1)}$$

where g_t is the $n_t \times 1$ vector of productivity growth for n_t constituents of a specific industry at time t. The main distinction between the two forms of learning arises from the sampling probability, $\text{Prob}(g_1, .., g_t | b, g_{c,1}, .., g_{c,t})$. Bayesian learning assigns equal weights to all past observations, so the sampling probability is given by

$$\text{Prob}(g_1, .., g_t | b, g_{c,1}, .., g_{c,t}) \propto \prod_{i=1}^{t} \exp\left(-\frac{(\bar{g}_i - b\bar{g}_{c,i}\mathbb{1}_{n_i})^\top(\bar{g}_i - b\bar{g}_{c,i}\mathbb{1}_{n_i})}{2\sigma^2}\right) \tag{E.2}$$

where $\bar{g}_{c,i}$ is the demeaned consumption growth in time i, $\bar{g}_i$ is the $n_i \times 1$ vector of the demeaned productivity growth, and $\mathbb{1}_{n_i}$ is the $n_i \times 1$ vector with all elements equal to one.

In contrast, fading memory allows differential weights to affect past observations depending on the time difference between observation and belief formation. Under this assumption, the sampling probability becomes

$$\text{Prob}(g_1, .., g_t | b, g_{c,1}, .., g_{c,t}) \propto \prod_{i=1}^{t} \left[\exp\left(-\frac{(\bar{g}_i - b\bar{g}_{c,i}\mathbb{1}_{n_i})^\top(\bar{g}_i - b\bar{g}_{c,i}\mathbb{1}_{n_i})}{2\sigma^2}\right)\right]^{(1-\nu)^{t-i}}. \tag{E.3}$$

where ν controls the speed of memory decay. While $\nu = 0$ is the Bayesian case, $\nu > 0$ implies fading memory. This equation is based on Nagel and Xu (2022).

This sampling probability (E.3) in combination with normally distributed prior beliefs with mean $m_{b,0}$ and standard deviation $\sigma_{b,0}$ leads to the posterior probability

$$\text{Prob}(g_1, .., g_t | b, g_{c,1}, .., g_{c,t}) \propto \exp\left(-\frac{(b - m_{b,t}^{\text{fade}})^2}{2(\sigma_{b,t}^{\text{fade}})^2}\right), \tag{E.4}$$

where the mean beliefs and precision are given by

$$m_{b,t}^{\text{fade}} = (\sigma_{b,t}^{\text{fade}})^2 \left[\frac{m_{b,0}}{\sigma_{b,0}^2} + \frac{1}{\sigma^2}\sum_{i=1}^{t}(1-\nu)^{t-i}\bar{g}_i^\top(\bar{g}_{c,i}\mathbb{1}_{n_i})\right] \tag{E.5}$$

$$\frac{1}{(\sigma_{b,t}^{\text{fade}})^2} = \frac{1}{\sigma_{b,0}^2} + \frac{1}{\sigma^2}\sum_{i=1}^{t}(1-\nu)^{t-i}\bar{g}_{c,i}^2 n_i. \tag{E.6}$$

Equations (E.5) and (E.6) imply that the belief distribution $\left(m_{b,t}^{\text{fade}}, \sigma_{b,t}^{\text{fade}}\right)$ are recursively updated as follows:

$$m_{b,t}^{\text{fade}} = \frac{\nu(\sigma_{b,t}^{\text{fade}})^2}{\sigma_{b,0}^2} m_{b,0} + \left(1 - \kappa_t(\sigma_{b,t}^{\text{fade}})^2 - \frac{\nu(\sigma_{b,t}^{\text{fade}})^2}{\sigma_{b,0}^2}\right) m_{b,t-1}^{\text{fade}} + \kappa_t(\sigma_{b,t}^{\text{fade}})^2 \hat{b}_t \tag{E.7}$$

$$\frac{1}{(\sigma_{b,t}^{\text{fade}})^2} = \frac{\nu}{\sigma_{b,0}^2} + \frac{1-\nu}{(\sigma_{b,t-1}^{\text{fade}})^2} + \kappa_t \tag{E.8}$$

$$\tag{E.9}$$

where $\kappa_t = n_t(\sigma_c\eta_t)^2/\sigma^2$ and $\hat{b}_t = \bar{g}_t^\top(\bar{g}_{c,t}\mathbb{1}_{n_t})/(n_t\sigma_c^2\eta_t^2)$. We note that when $\nu = 0$, the recursive equations (E.7) and (E.8) collapse to the Bayesian updating equations (3) and (4) of the main paper.

References

Malmendier, U., Nagel, S., 2016. Learning from inflation experiences. The Quarterly Journal of Economics 131, 53–87.

Nagel, S., Xu, Z., 2022. Asset pricing with fading memory. The Review of Financial Studies 35, 2190–2245.