2-1-2015

Intertemporal Information Acquisition and Investment Dynamics

Christian C. Opp
University of Pennsylvania

Follow this and additional works at: http://repository.upenn.edu/fnce_papers

Part of the Finance and Financial Management Commons

Recommended Citation

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/fnce_papers/10
For more information, please contact repository@pobox.upenn.edu.
Intertemporal Information Acquisition and Investment Dynamics

Abstract
This paper studies intertemporal information acquisition by agents that are rational Bayesian learners and that dynamically optimize over consumption, investment in capital, and investment in information. The model predicts that investors acquire more information in times when future capital productivity is expected to be high, the cost of capital is low, new technologies are expected to have a persistent impact on productivity, and the scalability of investments is expected to be high. My results shed light on the economic mechanisms behind various dynamic aspects of information production by the financial sector, such as the sources of variation in returns on information acquisition for investment banks or private equity funds.

Keywords
dynamic information acquisition, investment, asset markets, bayesian learning

Disciplines
Business | Finance and Financial Management

This working paper is available at ScholarlyCommons: http://repository.upenn.edu/fnce_papers/10
Intertemporal Information Acquisition and Investment Dynamics

Christian C. G. Opp*

February 2015

Abstract

The paper studies intertemporal information acquisition by agents that are rational Bayesian learners and that dynamically optimize over consumption, investment in capital, and investment in information. The model predicts that investors acquire more information in times when future capital productivity is expected to be high, the cost of capital is low, new technologies are expected to have a persistent impact on productivity, and the scalability of investments is expected to be high. My results shed light on the economic mechanisms behind various dynamic aspects of information production by the financial sector, such as the sources of variation in returns on information acquisition for investment banks or private equity funds.

JEL topic area codes: C61, D83, D91, D92, G24

1 Introduction

The investors’ conditional information sets drive the dynamics of both asset prices and investment. Which economic factors determine the evolution of the information sets? The asset pricing literature typically takes the information flow to investors as exogenous and uses various instruments to capture the dynamics of the information set.

In this paper, I endogenize the information set and explore the economic determinants of intertemporal information acquisition decisions. I develop a dynamic model in which investors are Bayesian learners who optimally choose how much to consume, how much to invest, and how much information to acquire. This model reveals various links between the investors’ efforts to learn

*The Wharton School, University of Pennsylvania. Email: opp@wharton.upenn.edu. This paper was developed during my time in the Chicago Booth Ph.D. program. I thank John Cochrane, John Heaton, Atif Mian, Thomas Sargent, Harald Uhlig, Pietro Veronesi, Arnold Zellner, and in particular my advisors Lars Hansen and Lubos Pastor, for outstanding support and feedback. This research was funded in part by the Best Foundation’s Arnold Zellner Doctoral Prize and by grants from the German National Academic Foundation and the German Academic Exchange Service (DAAD).
about new technologies and their conditional expectations on productivity, discount factors, and scale economies.

The model predicts that investors acquire more information when future capital productivity is expected to be high, expected returns are low, technology changes are expected to have a persistent impact on productivity, and future investment scalability, as for example influenced by liquidity, is expected to be high. Economic mechanisms that decrease the scalability of investment diminish the marginal gains from information acquisition since they restrict the flexibility to respond to news.

The results imply various predictions on the determinants of variation in knowledge acquisition by the financial service sector. Since human capital is typically the scarce factor in information production, compensation and hiring and firing decisions are expected to be linked to changes in financial institutions’ expectations about future productivity, persistence of technologies, and liquidity. Further, the theory rationalizes private equity and hedge funds’ preference for long-term capital commitments: lower dependence on future capital contributions gives the funds greater flexibility in intertemporal investment decisions. This increases returns on information acquisition. Firms that are more financially constrained are predicted to benefit less from information acquisition and thus will prefer to stay less informed. Further, private equity funds are predicted to stay less informed if they expect a less liquid IPO market and lower capital productivity in the future.

Generally, variants of the presented model might be useful to analyze differences in the information production technologies across investment firms and the corresponding determinants of the latent evolution of knowledge and investment skill. In addition, on a more aggregated level, the model could be used to analyze differences in the information acquisition technologies available to investors across countries, which could be useful for policy considerations: the framework links the characteristics of the information acquisition technology available to investors in a financial system to intertemporal investment and consumption allocations. This allows analyzing the impact of frictions to the intertemporal scalability of information acquisition, for example as imposed by labor market regulations, and changes in cost levels, for example as induced by accounting regulations.

Model Approach

To make the analysis tractable the presented model uses a reduced form representation of the equilibrium in the information market: an information acquisition technology subsumes various economic mechanisms that have an impact on the information acquisition cost faced by investors. A sequential joint consumption, investment, information acquisition problem is analyzed for a general time- and state-dependent information acquisition cost function and the class of additively separable preferences.

The resulting intertemporal optimality condition for information acquisition shows that the marginal return on uncertainty reduction satisfies - like all other
asset returns - the Euler equation, suggesting the interpretation of information as an asset. The paper explores the dynamic relation between the marginal return on this asset and technology changes.

**Literature**

Bayesian learning is an essential component of the presented framework. The literature on asset market dynamics caused by learning typically considers an exogenous evolution of information. Examples include Pastor and Veronesi [1] and Hansen [2]. A key distinction from this literature is the paper’s focus on active information acquisition decisions which make the information flow to agents endogenous.

Noisy rational expectations models in the spirit of Grossman and Stiglitz [10] are traditionally used to analyze information acquisition in finance (Verrecchia [11]; Barlevy and Veronesi [12]; Peress [13]; Veldkamp [14]). These models typically focus on aspects of asymmetric information and decentralization in static environments. When applied to the analysis of joint dynamics in information acquisition, investment, and consumption this type of setup quickly loses tractability despite restrictive assumptions about information cost structures, production technologies, and preferences. As shown later, the predictions of this paper arise through the analysis of a dynamic setup with richer specifications of technologies.

The paper shares with the recent literature on “rational inattention” (e.g., Sims [16], [17]; Luo [18]; Turmuhambetova [19]) that a dynamic setup is considered where information used by agents is endogenous. Yet, in contrast to the literature following Sims this paper does not model how an individual allocates her limited capacity to process available information. The literature following Sims is “motivated by the idea that information that is freely available to an individual may not be used, because of the individual’s limited information processing capacity” [20]. In contrast, this paper considers producers of information, like financial institutions that develop information that is not freely available at the time when it is used or sold. Financial institutions hire employees who use available information in combination with costly investigation to produce proprietary estimates of relevant investment measures. Sims explicitly excludes actions of “costly investigation” [20], such as interviewing experts on technologies and consumer surveys, from the activities where the constraints modeled in the rational inattention framework are of relevance. In this paper information acquisition cost refer to the full cost of developing information, such as the cost of labor, IT infrastructure, and physical assets. It is not the shadow price of a constraint to pay attention to information that an individual already

---

1 Other papers on robust control are e.g., Hansen, Sargent and Tallarini [3], Hansen and Sargent ([4], [5]) and Hansen, Sargent and Wang [6]. Learning is also a central concept to other parts of the recent finance literature such as Pastor and Veronesi [7] (learning about profitability), Lewellen and Shanken [8] (learning and asset pricing tests), Adrian and Franzoni [9] (learning about time-varying factor loadings).

2 Wang [15] considers a dynamic setup with asymmetric information but keeps information exogenous.
has at her disposal. This conceptually different focus leads to a different model approach. Agents are Bayesian learners that form rational expectations conditional on all information that is available to them.

Finally, Abel, Eberly, and Panageas consider a setup where inattentive agents update their information sporadically, rather than continuously. Optimal inattentive behavior is generated by the assumption that the consumer must pay a cost that is proportional to the portfolio’s contemporaneous value to observe the value of his investment portfolio. The consumer optimally checks his investment portfolio at equally spaced points in time, consuming from a riskless transactions account in the interim.

The paper proceeds as follows: the next Section builds a model of joint consumption, investment, and information acquisition decisions in a production economy with time varying technology risk. Intertemporal optimality conditions are presented and economic determinants of information acquisition by Bayesian learners are examined. The model’s predictions are discussed in reference to dynamic aspects of information production by the financial sector. Section 3 concludes.

2 Model

2.1 Production Technology

Consider a decision maker who may invest in a project with weakly decreasing returns to scale. Let \( F_t \) denote the decision maker’s information set at date \( t \), \( Y_t \) denote output at date \( t \) measured in terms of a single composite consumption good, and \( G_t (\cdot) \) be a strictly increasing, concave, and \( F_t \)-measurable function. Given risky capital \( k_t \) at date \( t \), output at date \( (t + 1) \) is determined by the following relation:

\[
Y_{t+1} = A_{t+1} G_t (k_t).
\]

In equation (1) the parameter \( A_{t+1} \) is stochastic conditional on the information set \( F_t \). Ex post, at date \( (t + 1) \), the decision maker may infer the actual realization of \( A_{t+1} \) based on the observed output \( Y_{t+1} \) and the choice of \( k_t \). Let \( \varepsilon_t = (\varepsilon_{a,t}, \varepsilon_{x,t}, \varepsilon_{\mu,t}, \varepsilon_{z,t})' \) be an i.i.d. Gaussian shock process with mean zero and identity covariance matrix. The productivity factor \( A_t \) is assumed to have the following exponential affine structure

\[
a_t \equiv \log A_t = \omega_a + \mu_t + \sigma_v \varepsilon_{a,t},
\]

where \( \omega_a \) sets the unconditional mean of \( a_t \), \( \mu_t \) is a stationary hidden state component with unconditional mean zero, and \( \varepsilon_{a,t} \) is an independent shock to productivity that prevents the full revelation of the hidden state \( \mu_t \) through the

---

4 Even though employees’ limits to pay attention to information they already have may influence their productivity this is likely not a central aspect for the analysis of financial institutions. Sims supports this view by stating that “sophisticated, continuously trading investors have uncertainty that is not freely available”. 
output realization \( Y_t \). \( \sigma_v \) is an exogenous scalar that determines the exposure of productivity to unpredictable noise \( \varepsilon_{a,t} \). In this paper a stationary economy is considered where the exogenous parameters \( \omega_a \) and \( \sigma_v \) are specified as time-invariant. Yet, the derivation in the Appendix accommodates time variation such that the analysis may be extended to study aspects like growth. Further, for notational simplicity the time subscript on the \( G_t(\cdot) \) function will usually be suppressed in the following exposition. For the stochastic simulation of the model that is presented below a Cobb-Douglas specification of the production function is considered which yields the following representation:

\[
Y_{t+1} = A_{t+1} k_{t+1}^{1-\eta} \quad \text{with} \quad 0 \leq \eta < 1.
\]  

In this specification the parameter \( \eta \) corresponds to the local curvature of the production technology.

**Production Technology Risk**

Empirical evidence suggests that the exposure of economies to technology risk exhibits time-variation (see e.g., Ng et al. [22]): during the internet or biotech "revolutions" (Pastor and Veronesi [1]) in the 1980s and 1990s economies like the US were exposed to substantial technological change and corresponding high levels of technological uncertainty. Investors were confronted with the task to predict the impact of the new technologies on productivity in the short, medium, and long run in order to make investment decisions. To do so they had to acquire information that allowed them to forecast changes in product markets, consumer behavior, and business prospects. The blooming of the profession of investment analysts during these times is anecdotal evidence for the increase in aggregate demand for information acquisition capacities. In contrast, in other times like in the years post 2000/2001 the capacities of analysts were markedly reduced, resulting in the well known layoff waves in the financial sector.

For a theory on dynamics in information acquisition by investors time-variation in technology risk thus seems to be an essential ingredient. In the model the notion of time varying exposures to technology risk is captured by the specification of the evolution of \( \mu_t \), the persistent component of log productivity \( a_t \). Technology risk therefore refers to shocks that also have an impact on productivity in future periods. It is this feature that makes “new technologies” a learning target for investors. In particular, \( \mu_t \) is assumed to follow a stationary Markov process that is mean reverting to zero:

\[
\mu_{t+1} = \phi_\mu \mu_t + \sigma_{\mu,t} \varepsilon_{\mu,t+1}, \quad 0 \leq \phi_\mu < 1.
\]  

The time-varying exposure to technology risk \( \sigma^2_{\mu,t} \) is governed by the Markov state \( z_t \) which evolves according to the following stationary process:

\[
z_{t+1} = \omega_z + \phi_z z_t + \sigma_z \varepsilon_{z,t+1}, \quad 0 \leq \phi_z < 1.
\]  

A \( F_t \)-measurable, positive-valued function \( \sigma^2_{\mu,t}(z_t) \) maps the state \( z_t \) into a risk exposure \( \sigma^2_{\mu,t} \). For the simulation results that are presented later in the paper
an EGARCH type of specification (Nelson [23]) of the function $\sigma_{\mu,t}^2 (z_t)$ is used:

$$\sigma_{\mu,t}^2 = \varphi \sigma \exp \{ z_t \}, \quad \varphi > 0. \quad (6)$$

### 2.2 Information Acquisition Technology

In reality, a variety of economic factors have an impact on the information acquisition technology available to investors. Essentially the financial system as a whole determines the cost of acquiring information on projects in the economy: laws and regulations, the accounting system with institutions like the SEC, the development of the banking system and financial intermediaries, availability of human capital, and the micro-structure of stock markets and trading platforms, to mention only a short list of potentially relevant factors.

A broad finance and accounting literature suggests that a system of efficient disclosure rules, financial intermediaries (Diamond [24]), and financial contracts (Townsend [25]) reduces information asymmetries and mitigates the duplication of effort. Law systems influence corporate governance and investors’ facility to monitor a company’s operations and outcomes in timely fashion (La Porta et al. [26]; D’Avolio et al. [27]; Dyck et al. [28]).

Trading technologies and the structure of information markets influence the dissemination and aggregation of information held by market participants and thus the cost of information acquisition faced by investors (Diamond and Verrecchia [29]; Hellwig [30]; Admati [31]; Admati and Pfleiderer [32], [33], [34]; Veldkamp [14]).

In order not to get lost in the variety of these mechanisms and to maintain a focus on joint dynamics in investment, consumption, and information acquisition this paper proposes a reduced form approach to characterize the information acquisition technology available to investors. The technology, specified by a cost function over information precision, subsumes various economic mechanisms that determine the costliness of information, including effects induced by informational externalities. The cost structure is the outcome of equilibria in other markets that are of relevance for information acquisition, or in short by the “information market.” The decision maker takes this cost structure as exogenously given. Quantity decisions are assumed to be not publicly observable. The price of capital is equal to one in terms of the composite consumption good and does not signal additional information.

In terms of the model the decision maker may influence the precision of a signal $s_t$ that she obtains jointly with the production output at the end of each period. The signal structure is specified as follows:

$$s_{t+1} = \mu_{t+1} + \frac{\sigma_{\mu}}{\sqrt{t}} \varepsilon_{s,t+1}. \quad (7)$$

The strictly positive and continuous decision variable $\iota_t$ controls information acquisition by scaling the noise variance of the signal $s_{t+1}$. Obtaining one signal

---

4 Laws also influence investors’ protection against expropriation by insiders and corporate governance control rights.
with noise variance $\sigma^2_v/\iota_t$ is informationally equivalent to obtaining $\iota_t$ separate, independent signals with noise variance $\sigma^2_v$; conditional uncertainty is reduced by the same amount. Since the parameter $\sigma_v$ is common to the specifications of $s_{t+1}$ and $a_{t+1}$ the scaling implies that choosing $\iota_t = 1$ corresponds to the acquisition of information that has the same signal-to-noise ratio as the period’s log-productivity realization $a_{t+1}$.

The cost of acquiring a signal $s_{t+1}$ with noise variance $\sigma^2_v/\iota_t$ is specified by the function $n_t(\iota_t)$. The model is analyzed for a general time and state dependent, strictly increasing and convex cost function which is required to satisfy $n'_t(\iota_t)|_{\iota_t=0} = 0$, for all $t$, to ensure strictly positive choices of $\iota$. The convexity of $n_t(\iota_t)$ reflects the notion that acquiring an additional independent signal becomes more costly the more information is acquired in a given period. For the simulation of the model, the following functional form will be considered:

$$n_t(\iota_t) = \chi \kappa + 1 \cdot \iota_t^{\kappa+1}, \text{ with } \kappa \geq 0 \text{ and } \chi > 0. \quad (8)$$

The specification features two properties of an information acquisition technology: the marginal cost of information acquisition and the costliness of varying learning effort intertemporally. The parameter $\chi$ shifts the marginal cost, the parameter $\kappa$ influences the local curvature of $n_t(\iota_t)$.

**Bayesian Updating**

Due to the conditional knowledge of the Markov state $z_t$ which governs the risk exposures $\sigma^2_{\mu,t}$ a version of Kalman filtering with time varying signal precision is applicable. As a Bayesian learner, the agent updates his beliefs about the hidden productivity state based on observed realizations of the random variables $a_t$ and $s_t$. Let $(a^t, s^t)$ denote the whole history of signals that the decision maker obtained up to date $t$, i.e.,

$$(a^t, s^t) \equiv \{(a_t, s_t), (a_{t-1}, s_{t-1}), (a_{t-2}, s_{t-2}), \ldots\},$$

and define the conditional mean $\hat{\mu}_t \equiv E[\mu_{t+1}|a^t, s^t]$ and the conditional variance $\hat{\sigma}^2_t \equiv Var[\mu_{t+1}|a^t, s^t]$. Given the initial prior distribution of the hidden productivity state is normal, $\mu_t \sim N(\hat{\mu}_0, \hat{\sigma}^2_0)$, the posterior density at date $t$ is also normal, $\mu_{t+1} \sim N(\hat{\mu}_t, \hat{\sigma}^2_t)$, where the conditional mean $\hat{\mu}_t$ and the conditional variance $\hat{\sigma}^2_t$ follow from the Kalman filter recursions:

$$\hat{\mu}_t = \phi_{\mu} \frac{a_t - \omega_{\mu,t-1} + \iota_{t-1}s_t + \sigma^2_{\mu,t-1}\hat{\mu}_{t-1}}{1 + \iota_{t-1} + \sigma^2_{\mu,t-1}}$$

$$\hat{\sigma}^2_t \equiv Var[\mu_{t+1}|a^t, s^t] = \sigma^2_{\mu,t} + \hat{\sigma}^2_{p,t}, \quad (9a)$$

It is required that the state that governs variation in the functional form of $n_t(\iota_t)$ is exogenous to the decision maker.
where \( \hat{\sigma}_{p,t}^2 \) denotes the pre-shock conditional variance at time \( t \) which is defined as follows:

\[
\hat{\sigma}_{p,t}^2 (\hat{\sigma}_{t-1,t}^2) \equiv \phi_\mu \frac{\sigma_v^2}{1 + \iota_{t-1} + \frac{\sigma_v^2}{\hat{\sigma}_{t-1}^2}}.
\]  

(10)

### 2.3 Decision Problem

The decision maker solves a sequential joint consumption-, investment-, information acquisition problem. Choices are ranked by von Neumann Morgenstern preferences. Let \( \beta_t \in (0, 1) \) be a \( F_t \)-measurable discount factor for the time period from \( t \) to \( (t + 1) \) and \( F (c_t) \) be a strictly increasing and concave one-period utility function. Further, let \( y_t \equiv \log Y_t \) denote log-output. Given an initial state vector \( x_0 \equiv (y_0, \hat{\mu}_0, \hat{\sigma}_{p,0}^2, z_0) \), consider the sequential problem to maximize

\[
F (c_t) + E_t \sum_{j=1}^{\infty} \left( \prod_{i=0}^{j-1} \beta_{t+i} \right) F (c_{t+j})
\]

subject to the budget restriction

\[
c_t = Y_t - k_t - n_t (\iota_t)
\]

and the transition laws and technologies specified in equations (1) - (10). Associated with this problem is the Bellman equation

\[
V (x_t) = \max_{u_t = (k_t, \iota_t)} \{ F (c_t) + \beta_t E [V (x_{t+1}) | x_t, u_t] \}
\]

(13)

with the vector of state variables \( x_t \equiv (y_t, \hat{\mu}_t, \hat{\sigma}_{p,t}^2, z_t)^\prime \), and the vector of controls \( u_t \equiv (k_t, \iota_t)^\prime \), and where \( V (x_t) \) denotes the optimal value of the objective function (11) starting from state \( x_t \). \( E [\cdot | x_t, u_t] \) denotes the mathematical expectation given the vectors \( (x_t, u_t) \) which yield a sufficient statistic for the conditional distribution of the next period state \( x_{t+1} \). While \( \hat{\mu}_t \) and \( \hat{\sigma}_{p,t}^2 \) characterize the conditional distribution of the hidden state \( \mu_{t+1} \), the observable Markov state \( z_t \) determines the distribution of the next period technology risk exposure \( \sigma_z^2 (z_{t+1}) \). The state vector transition equations are given below.

\[
\begin{align*}
y_{t+1} &= a_{t+1} + \log G_t (k_t) \\
\hat{\mu}_{t+1} &= \phi_\mu \frac{\sigma_e^2}{1 + \iota_t + \frac{\sigma_e^2}{\hat{\sigma}_{t-1}^2}} \hat{\mu}_t \\
\hat{\sigma}_{p,t+1}^2 &= \phi_\mu^2 \frac{\sigma_v^2}{1 + \iota_t + \frac{\sigma_v^2}{\hat{\sigma}_{t-1}^2}} \\
z_{t+1} &= \omega_z + \phi_z \hat{z}_t + \sigma_z \varepsilon_{z,t+1}
\end{align*}
\]

(14)
2.4 Optimality Conditions

The envelope condition and the first order necessary condition for capital $k_t$ yield the familiar Euler equation

$$E \left[ A_{t+1} G' (k_t) m_{t+1} \mid x_t, u_t \right] = 1,$$

(15)

where $m_{t+1} = \frac{\beta_t F'(c_{t+1})}{\hat{\sigma}_{p,t+1}^2}$ denotes the stochastic discount factor and $A_{t+1} G' (k_t)$ the marginal gross return on capital. Note that information acquisition decisions play a central role in the Euler equation: by controlling information acquisition $\iota_t$, the investor influences the evolution of the conditional information set and thus the conditional expectations operator in the Euler equation (15). Dynamics in asset prices are thus directly linked to decisions on $\iota_t$ and their economic determinants. This is a key distinction from the existing literature that analyzes dynamics induced by Bayesian learning but takes the information flow as exogenous. In reference to the language used in empirical asset pricing, the set of signals obtained by the investor may be viewed as an “endogenous instrument” that summarizes conditional information.

Optimization generates restrictions on intertemporal information acquisition and thus on the information flow to investors. Let $\lambda$ denote Lebesgue measure and $f (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1} \mid x_t, u_t)$ the joint conditional probability density function of the vector $(a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})'$. The first order necessary condition for information acquisition $\iota_t$ may be written as follows:

$$\frac{n_t' (\iota_t)}{\beta_t} F' (c_t) = \int \left( \frac{\partial \hat{\mu}_{t+1}}{\partial \iota_t} \frac{\partial V (x_{t+1})}{\partial \mu_{t+1}} + \frac{\partial \hat{\sigma}_{p,t+1}^2}{\partial \iota_t} \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}_{p,t+1}^2} \right) \cdot f (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1} \mid x_t, u_t) d\lambda (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1}) + \int V (x_{t+1}) \frac{\partial f (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1} \mid x_t, u_t)}{\partial \iota_t} d\lambda (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})$$

(16)

where $\hat{\mu}_{t+1} = \hat{\mu}_{t+1} (a_{t+1}, s_{t+1}, \iota_t, \hat{\sigma}_{p,t+1}^2, \iota_t)$ and $\hat{\sigma}_{p,t+1}^2 = \hat{\sigma}_{p,t+1}^2 (\hat{\sigma}_{p,t+1}^2, \iota_t)$. The following Proposition gives the associated necessary intertemporal optimality condition.

**Proposition 1.** The decision problem associated with the Bellman equation [13] implies the following necessary intertemporal condition for optimal information acquisition $\iota_t$:

$$E_t \left[ m_{t+1} \frac{\Pi_{t+1} + \delta_{t+1} M C_{t+1} (\hat{\sigma}_{p,t+1}^2) }{M C_t (\hat{\sigma}_{p,t+1}^2)} \right] = 1$$

(17)

where $m_{t+1} = \frac{\beta_t F'(c_{t+1})}{F'(c_t)}$ denotes the stochastic discount factor and the following
definitions apply

\[ \Pi_{t+1} = \frac{1}{2} \frac{\partial^2 V(x_{t+1})}{\partial \hat{\mu}_{t+1}^2} - \frac{\partial^2 V^{RS}(x_{t+1}, u_{t+1})}{\partial \mu_{t+1}^2} F'(c_{t+1}) \]

\[ = \frac{1}{2} \frac{\partial^2 V^{RS}(x_{t+1}, u_{t+1})}{\partial \mu_{t+1}^2} \frac{\partial k_{t+1}(x_{t+1})}{\partial \mu_{t+1}} + \frac{\partial^2 V^{RS}(x_{t+1}, u_{t+1})}{\partial \mu_{t+1}^2} \frac{\partial k_{t+1}(x_{t+1})}{\partial \mu_{t+1}} \]

\[ MC_t (\hat{\sigma}_{p,t+1}^2) \]

\[ = n_t' (\mu_t) \left\{ - \frac{\partial \hat{\sigma}_{p,t+1}^2 (\hat{\sigma}_{t+1}^2, t)}{\partial t} \right\} \]

\[ = - \frac{\phi_\mu^2 \sigma_{v,t}^2}{\left(1 + \mu_t + \frac{\sigma_{v,t}^2}{\sigma_t^2}\right)^2} \]

\[ \delta_{t+1} = \frac{\partial \hat{\sigma}_{p,t+2}^2 (\hat{\sigma}_{t+1}^2, t, t+1)}{\partial \hat{\sigma}_{t+1}^2} = \frac{\phi_\mu^2}{\left(1 + (1 + \mu_{t+1}) \frac{\hat{\sigma}_{t+1}^2}{\sigma_{t+1}^2}\right)^2} \]

and where \( V^{RS} (x_{t+1}, u_{t+1}) \) denotes the right side of the Bellman equation \[13\] evaluated at the optimal policy rule \( u(x_{t+1}) \).

**Proof.** See Appendix. \qed

**Discussion of Proposition 1**

Equation (17) indicates that at the optimum, the marginal return on uncertainty reduction satisfies an Euler equation: current marginal cost of reducing uncertainty, denoted by \( MC_t (\hat{\sigma}_{p,t+1}^2) \), are compensated by a future, stochastic benefit which may be decomposed into two components: \( \Pi_{t+1} \), which is accruing at time \((t+1)\), and subsequent benefits of value \( \delta_{t+1} MC_t (\hat{\sigma}_{p,t+2}^2) \), where \( \delta_{t+1} \) captures the notion of depreciation of information. The structure of this marginal return is similar to the one of a return on capital where the next period benefit consists of a dividend payment and a remaining depreciated capital value. The representation suggests that marginal uncertainty reduction induced by information acquisition may be viewed as an asset that generates a stream of benefits in future periods. In the following, the terms in equation (17) are discussed in more detail.

\( \Pi_{t+1} \): Marginal Next Period Benefit of Uncertainty Reduction

\( \Pi_{t+1} \), the marginal next period benefit of uncertainty reduction, is generated by state-contingent adjustments to tomorrow’s controls \( u(x_{t+1}) \) in reaction to a more precise productivity estimate. Aversion against uncertainty about the next period productivity estimate \( \hat{\mu}_{t+1} \), as represented by the curvature of the value function \( (\partial^2 V (x_{t+1}) / \partial \mu_{t+1}^2) \), is alleviated by additional information acquisition since next period’s decisions optimally take into account additional information. SubSection 2.5 considers an approximation of \( \Pi_{t+1} \) which sheds more light on the economic determinants of this term.
Marginal Cost of Uncertainty Reduction

$MC_t \left( \hat{\sigma}_{p,t}^2 \right)$ denotes the marginal cost of reducing next period’s conditional uncertainty about the hidden productivity state $\mu_{t+1}$. The subscript $t$ indicates that the marginal cost function $MC_t \left( \hat{\sigma}_{p,t}^2 \right)$ is $\mathcal{F}_t$-measurable, it depends on the uncertainty level $\hat{\sigma}_{p,t}^2$ at date $t$. The marginal cost are influenced by (a) the marginal cost of information $n_t' (\iota_t)$ as determined by the information acquisition technology and (b) the marginal impact of information on conditional uncertainty $(\partial \hat{\sigma}_{p,t+1}^2 / \partial \iota_t)$ which is determined by the Ricatti-difference equation resulting from Kalman filtering. Convexity of $n_t(\iota_t)$ and strict convexity of $\hat{\sigma}_{p,t+1}^2$ ensure that the function $MC_t \left( \hat{\sigma}_{p,t}^2 \right)$ is strictly declining in $\hat{\sigma}_{p,t+1}^2$ over its domain $(0; \hat{\sigma}_{p,t}^2)$. In other words, the lower the choice of next period’s pre-shock conditional variance $\hat{\sigma}_{p,t+1}^2$ the more costly is a further marginal reduction. The marginal cost are high when the current level of uncertainty $\hat{\sigma}_{p,t}^2$ is already relatively low and/or a lot of information is acquired in this period. $MC_t \left( \hat{\sigma}_{p,t+1}^2 \right)$ goes to infinity as $\hat{\sigma}_{p,t+1}^2$ approaches zero.

Depreciation of Information

The $\mathcal{F}_{t+1}$-measurable parameter $\delta_{t+1}$ measures by how much a marginal decline in risk at date $(t+1)$ reduces risk at date $(t+2)$, i.e., it quantifies an intertemporal informational externality. Information on the hidden productivity state at time $t$ also contains information on the hidden state at time $(t+1)$. Yet, the predictive precision is reduced $(0 < \delta_{t+1} < 1)$. $\delta_{t+1}$ has the interpretation of a depreciation factor that measures the speed at which information becomes irrelevant. In equation (17) $MC_{t+1} \left( \hat{\sigma}_{p,t+2}^2 \right)$ is therefore depreciated by the factor $\delta_{t+1}$. Obsolescence of information is caused by technological change and thus influenced by the process parameters that determine the persistence of productivity. Higher values of $\phi_\mu$ increase the persistence of the hidden productivity state $\mu_t$ and therefore increase $\delta_{t+1}$. The parameter $\phi_z$ which influences the persistence of risk exposures has a more ambivalent impact on $\delta_{t+1}$: conditional on a high level of technology risk $\sigma_{\mu}^2 (z_t)$, high values of $\phi_z$ tend to diminish $\delta_{t+1}$ since higher future technology shocks tend to diminish the persistence of $\mu_t$. On the other hand, conditional on low technology risk at time $t$, higher values of $\phi_z$ tend to increase $\delta_{t+1}$.

Stochastic Discount Factor

Just like the return on any other asset, the return on information is discounted by the stochastic discount factor $m_{t+1}$. Conversely, a given discount factor implies restrictions on the marginal returns on information acquisition. Overall, equation (17) ensures that the decision maker sets today’s marginal utility sacrifice equal to the expected utility gain arising from future state contingent decision adjustments in reaction to a more precise productivity estimate.
2.5 The Marginal Benefit \( \Pi_{t+1} \)

The following two Propositions analyze the term \( \Pi_{t+1} \) in more detail.

**Proposition 2.** The addends that enter \( \Pi_{t+1} \) may be written as follows:

\[
\frac{\partial^2 V^{RS}(x_{t+1}, \mathbf{u}_{t+1})}{\partial \mu_{t+1} \partial t_{t+1}} \left. \frac{\partial k_{t+1}(x_{t+1})}{\partial t_{t+1}} \right|_{F'(c_{t+1})} = \text{Cov}_{t+1} \left[ s_{t+2}, A_{t+2} G'(k_{t+1}) m_{t+2} \right] \frac{\partial k_{t+1}(x_{t+1})}{\partial \mu_{t+1}}
\]

\[
\frac{\partial^2 V^{RS}(x_{t+1}, \mathbf{u}_{t+1})}{\partial \mu_{t+1} \partial t_{t+1}} \left. \frac{\partial k_{t+1}(x_{t+1})}{\partial t_{t+1}} \right|_{F'(c_{t+1})} = - \frac{\phi^2_{t} \beta_{t+1} \sigma_{e_{t+1}}^2}{F'(c_{t+1})} \left( 1 + \mu_{t+1} + \sigma_{e_{t+1}}^2 \frac{\partial \mu_{t+1}}{\partial \mu_{t+1}} \right) \cdot \frac{\partial t_{t+1} (x_{t+1})}{\partial \mu_{t+1}}.
\]

\[
\left\{ \begin{array}{l}
E_{t+1} \left[ \frac{\partial V(x_{t+2})}{\partial \sigma^2_{p,t+2}} - \frac{1}{2} \frac{\partial^2 V(x_{t+2})}{\partial \sigma^2_{p,t+2}} \right] \\
E_{t+1} \left[ \frac{\partial V(x_{t+2})}{\partial \sigma^2_{e,t+2}} - \frac{1}{2} \frac{\partial^2 V(x_{t+2})}{\partial \sigma^2_{e,t+2}} \right]
\end{array} \right. +
\]

\[
E_{t+1} \left[ \phi_{t} \cdot \partial - \frac{\partial V(x_{t+2})}{\partial \sigma^2_{e,t+2}} - \frac{1}{2} \frac{\partial^2 V(x_{t+2})}{\partial \sigma^2_{e,t+2}} \right]
\]

For small \( \phi_{t} \), and given a regularity condition is satisfied\(^6\) the second addend is, to a first order approximation, equal to zero, i.e.,

\[
\frac{\partial^2 V^{RS}(x_{t+1}, \mathbf{u}_{t+1})}{\partial \mu_{t+1} \partial t_{t+1}} \left. \frac{\partial k_{t+1}(x_{t+1})}{\partial t_{t+1}} \right|_{F'(c_{t+1})} \approx \phi^2_{t} b(x_{t+1}, \phi_{t}) \Big|_{\phi_{t}=0} + \frac{d \left( \phi^2_{t} b(x_{t+1}, \phi_{t}) \right)}{d \phi_{t}} \Big|_{\phi_{t}=0} \phi_{t} = 0.
\]

**Proof.** See Appendix. \( \square \)

\(^6\)Define the function

\[
b(x_{t}, \phi_{t}) = - \frac{\sigma^2_{u,t} \beta_{t}}{1 + \mu_{t} (x_{t}, \phi_{t}) + \sigma^2_{u,t}}
\]

\[
E_{t} \left[ \phi_{t} \cdot \partial - \frac{\partial V(x_{t+1}(\phi_{t}), \phi_{t})}{\partial \sigma^2_{e,t+1}} - \frac{1}{2} \frac{\partial^2 V(x_{t+1}(\phi_{t}), \phi_{t})}{\partial \sigma^2_{e,t+1}} \right] \cdot \frac{\partial t_{t+1} (x_{t+1})}{\partial \mu_{t+1}}
\]

We require that

\[
\left\{ \begin{array}{l}
\left| \phi_{t} b(x_{t}, \phi_{t}) \right|_{\phi_{t}=0} = 0 \\
\left| \phi^2_{t} \frac{\partial b(x_{t}, \phi_{t})}{\partial \phi_{t}} \right|_{\phi_{t}=0} = 0.
\end{array} \right.
\]

12
Proposition 3. Given that the period return function $F(c_t)$ is linear, $G(k_t)$ is strictly concave, and given the regularity condition stated in Proposition 2 is satisfied, the following first order approximation of $Π_{t+1}$ around $φ_μ = 0$ holds:

$$Π_{t+1}(x_{t+1}, φ_μ) ≈ Π_{t+1}(x_{t+1}, φ_μ)|_{φ_μ=0} + \frac{∂Π_{t+1}(x_{t+1}, φ_μ)}{∂φ_μ}|_{φ_μ=0} φ_μ$$

$$= \frac{k_{t+1}(x_{t+1})}{2\eta_{t+1}(x_{t+1})},$$

where

$$η_{t+1}(x_{t+1}) ≡ -\frac{G''(k_{t+1}(x_{t+1}))}{G'(k_{t+1}(x_{t+1}))} k_{t+1}(x_{t+1}).$$

If, in addition, $G(k_t) = k_t^{1-η}$, with $0 < η < 1$, we obtain

$$Π_{t+1}(x_{t+1}) ≈ \frac{((1-η)E_{t+1}[A_{t+2}] \beta_{t+1})^\frac{1}{η}}{2η}.$$

Proof. See Appendix.

Discussion

Proposition 2 shows that, to a first order approximation around $φ_μ = 0$, the state contingent adjustments to the control $ι_{t+1}$ in response to additional information have a negligible impact on the marginal benefit of information acquisition.

In Proposition 3 the setup is specialized to the case where the curvature from the period return function is set to zero. This case turns out to be instructive since the results become more tractable. Linear utility implies that the decision maker essentially maximizes discounted cash flows, where $\{c_t\}$ represents the dividend process that the project generates. The assumption of exogenous discount factors $β$ shuts off the potential incentive channel to reduce future cost of capital, but substantially increases the tractability of the analysis. The investment project is assumed to exhibit strictly declining returns to scale, i.e., $G''(k_t) < 0$. The agent decides when and how much capital to place in the project and when and how much resources to commit to research activities.

This setup captures features of the decision problem of investment banks and other financial institutions like private equity funds which, over time, decide on the amount of resources spent on investment research. In practice, an institution’s capacity to acquire information on investments will largely depend on the quality and quantity of its staff. The costliness of intertemporal changes to research capacities are influenced by the labor market and frictions induced by laws and regulations. Such aspects may be captured by the curvature of the cost function $n(ι)$. Further, recall that $ι$ denotes the amount of independent information acquired. With an increasing number of hired analysts the marginal analyst likely generates less additional independent information. This aspect also contributes to the convexity of the cost function $n(ι)$. 
Proposition 3 indicates that the marginal benefit $\Pi_{t+1}$ is not level independent. It depends on the scale of investment $k_{t+1}$. This paper considers the case of stationary productivity. In the case of growth, adjustments to the information acquisition technology would have to capture the increasing cost of obtaining comparable information levels for larger projects.

Uncertainty reduction yields a more precise conditional productivity estimate in each future state. This in turn allows sharper variations in the decisions in reaction to different estimates in the various future states of the world. For example, in the state of a high conditional productivity estimate, more precision might encourage the decision maker to reduce consumption more in favor of investment. Yet, curvature on the production function, as characterized by the elasticity $\eta_t$, which diminishes the variability of investment, reduces the marginal benefit of information acquisition. The elasticity scales down the benefit since it reduces the decision maker’s flexibility to adjust future controls in response to different signals in different states of the world. The following comparative statics summarize the results from Proposition 3.

**Comparative Statics (in Ref. to Proposition 3)**

The benefit of information acquisition is higher in times:

1. when productivity is expected to be high
2. when discount factors are high (less discounting)
3. when the scalability of investments is high
4. when technology changes are expected to have a persistent impact on productivity.

The simulations of the model presented below mirror these results. The plots displayed in Figure I show changes in average information acquisition as induced by changes in different exogenous parameters of the model. Details of the simulation are given in the description below Figure I.

Higher marginal cost of information acquisition, as influenced by the parameter $\chi$, decrease information acquisition. A higher value of $\eta$ decreases the scalability of investment and thus reduces information acquisition $\iota$. High persistence in productivity, as captured by the parameter $\phi_\mu$, increases the future value of information about the current state of productivity and therefore encourages information acquisition. The parameter $\omega_z$ changes the unconditional level of technology risk $\sigma^2_\mu$: higher exogenous risk exposures increase the incentive to obtain information (Note that due to the exponential specification of $\sigma^2_\mu$ the case of $\omega_z = 0$ does not correspond to the absence of technology risk.). Finally, the parameter $\omega_a$ varies the unconditional productivity level: higher productivity increases capital investment $k$ and thus the incentive to acquire information.
FIGURE I: The plots show percentage changes in average information acquisition $\iota$, relative to the level under the base parameter setup listed in Table 1. Variation is caused by changes in the exogenous parameters indicated on the horizontal axes. All other parameters are fixed at their base levels. An exception is the plot that considers variation in $\eta$: here the parameter $\omega_a$ is adjusted simultaneously to keep the deterministic steady state output level fixed. Without this adjustment the effect goes in the same direction and is even more pronounced. The plots are based on a stochastic simulation of the model under the assumptions made in Proposition 3. A second order Taylor approximation of the decision and transition functions of the model are computed. The method uses a perturbation approach (Judd [35]) for computing the quadratic approximation of the policy function. The base parameter setup for the simulations is provided in Table I.
<table>
<thead>
<tr>
<th>Base Parameter Setup</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(\phi_\mu)</td>
<td>0.5000</td>
<td>(\beta)</td>
</tr>
<tr>
<td>(\eta)</td>
<td>0.5000</td>
<td>(\sigma_z)</td>
</tr>
<tr>
<td>(\omega_a)</td>
<td>0.3722</td>
<td>(\sigma_v)</td>
</tr>
<tr>
<td>(\chi)</td>
<td>0.0050</td>
<td>(\phi_z)</td>
</tr>
<tr>
<td>(\kappa)</td>
<td>4.0000</td>
<td>(\omega_z)</td>
</tr>
<tr>
<td>(\varphi_\sigma)</td>
<td>0.0030</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Simulated Mean Values</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(Y)</td>
<td>1.0866</td>
<td>(a) 0.3780</td>
</tr>
<tr>
<td>(c)</td>
<td>0.5683</td>
<td>(s) 0.0016</td>
</tr>
<tr>
<td>(\kappa)</td>
<td>0.5162</td>
<td>(\hat{\mu}) 0.0033</td>
</tr>
<tr>
<td>(\iota)</td>
<td>0.6109</td>
<td>(\sigma_p^2) 0.0038</td>
</tr>
<tr>
<td>(\hat{\sigma}^2)</td>
<td>0.0094</td>
<td>(z) 0.0059</td>
</tr>
</tbody>
</table>

Table 1: The table lists the base parameters and simulated mean values of endogenous variables from the solution (see additional details in the caption of Figure I).

**Interpretation of the Years Pre and Post 2000/2001**

The results may shed some light on the substantial employment fluctuations at major investment banks observed in the years pre and post 2000/2001. Traditionally, especially Anglo-Saxon investment houses have a reputation for substantial employment fluctuations. In 1996, for example, headcount at Goldman Sachs was 6000; by 2001, it was 23,000 and by 2003 it was down to 19,500. Merrill Lynch increased the number of employees from 46,000 in 1995 to 72,000 in 2000, and then made 24,000 redundancies globally in two years from mid 2001 to mid 2003. As shown in Figure II, these examples were no exceptions — other top investment banks saw similar employment fluctuations.
FIGURE II: Employment fluctuations at major investment banks. This figure plots changes in the number of employees at major investment banks relative to their level in 1996.

In the late 1990s conditional productivity estimates were presumably good. Investors were likely to expect that the new internet technologies would have a persistent impact on future productivity, making them a valuable object to learn about. At the same time, according to the analysis, low expected returns would have further increased the incentive to acquire information.

In contrast, post 2000/2001, productivity forecasts were revised downwards markedly (compare Figure 6 in Pastor and Veronesi [36] on profitability). Markets exhibited low liquidity which increased the cost of adjusting investment levels in reaction to information. Further, accounting scandals were essentially deteriorating the financial system’s information acquisition technology; they effectively increased the cost of information acquisition faced by investors. The cheap signal from financial records turned out to be noisier than previously thought. Finally, risk prices increased quite substantially.

Overall, these stylized facts seem to match the predictions on the dynamic links between information acquisition and the asset market fairly well. Labor is a costly overhead for financial institutions, frictions to changing the size of the work force may be quite substantial. The results suggests that hiring and firing
decisions by the financial sector might be an interesting signal of the institutions’ conditional expectations on productivity, expected returns, and liquidity.

**Long Term Investments, Financial Constraints, and Liquidity**

The model also sheds some light on the fact that private equity funds and hedge funds benefit more from investment research if clients commit capital for a longer time horizon. Long-term capital commitments reduce outside restrictions on the funds’ intertemporal investment allocation and thus increase the value of marginal knowledge. The results also suggest that private equity funds’ information acquisition on start-ups is more valuable when liquid IPO markets, high productivity, and low cost of capital are anticipated in the future.

Generally, restrictions on future investment decisions as for example imposed by required cash outflows are predicted to decrease an institution’s investments in information: it binds the intertemporal investment policy and diminishes the flexibility to respond to marginal information. Firms that are financially constrained and thus face limitations to the scalability of their investments are likely to invest less in research on the productivity of available projects.

**2.6 Human Capital and Information Acquisition**

In this Section I consider a slightly adjusted setup where the agent has some human capital or skill to acquire information. The level of skill is subject to depreciation and there are adjustment cost. The level of human capital will be denoted by $h$. To ensure tractability, the hidden productivity state $\mu_t$ is now assumed to be independently distributed over time such that signals have value in exactly one period. This setup is convenient since it will allow us to write the marginal return on information acquisition in terms of known functions of consumption, information acquisition, and the level of human capital, without the need for an approximation. Similar to before, we consider an affine specification for $a_{t+1}$:

$$a_{t+1} = \omega_{a,t} + \mu_t + \sigma_a \varepsilon_{a,t+1}. \quad (19)$$

Notice that there is a slight twist in the timing relative to the setup used so far: $a_{t+1}$ is now a function of $\mu_t$, not of $\mu_{t+1}$. Since the signal structure will be left unchanged (as in equation (7)) this timing change is necessary for the signal to have value under the assumption that $\mu_t$ is independently distributed over time. Specifically, $\mu_t$ is now assumed to be normal with zero mean and variance $\sigma^2_{\mu,t-1}$, where, as before $\sigma^2_{\mu,t-1}$, is specified as a function of the Markov state $z_{t-1}$:

$$\mu_t = \sigma^2_{\mu,t-1} \varepsilon_{\mu,t}. \quad (20)$$

To allow for persistent changes in conditional expectations about productivity $\omega_{a,t}$ is now assumed to follow a stationary Markov process:

$$\omega_{a,t+1} = \bar{\omega} + \phi_\omega \omega_{a,t} + \sigma_\omega \varepsilon_{\omega,r+1}. \quad (21)$$
In this Section, skill to acquire information is determined by past information acquisition activity. This assumption is motivated by the notion of learning on the job: the more information the agent acquired in the past the more proficient she has become, and the fewer resources she needs to exert to acquire new information today. To capture this notion in the simplest form, the level of human capital $h_t$ will be set equal to last period’s information acquisition level $\iota_{t-1}$, depreciated by a factor $\delta_h$ that accounts for the speed with which this type of human capital becomes obsolescent, i.e., it characterizes the persistence in skill:

$$h_t = \delta_h \iota_{t-1} \quad \text{with } 0 \leq \delta_h < 1.$$  \hfill (22)

For simplicity, I consider a quadratic information acquisition cost function that features the notion that higher levels of human capital $h_t$ allow the agent to acquire information at lower cost:

$$n(\iota_t, h_t) = \chi \frac{(\iota_t - h_t)^2}{2} \quad \text{with } \chi > 0.$$  \hfill (23)

The budget constraint is now given by

$$Y_t = k_t + n(\iota_t, h_t) + c_t.$$  \hfill (24)

The level of human capital $h_t$ and the conditional log productivity parameter $\omega_{a,t}$ yield two additional state variables. The state may thus be summarized by the vector $x_t = (y_t, h_t, \omega_{a,t}, \hat{\mu}_t, \hat{\sigma}_{t}^2, z_t)'$. The corresponding state vector transition equations are given by:

$$
\begin{align*}
y_{t+1} &= a_{t+1} + \log \{G(k_t)\} \\
h_{t+1} &= \delta_h t_t \\
\omega_{a,t+1} &= \bar{\omega} + \phi_{a} \bar{\omega}_{a,t} + \sigma_{a} \bar{\omega}_{a,t+1} \\
\hat{\mu}_{t+1} &= \frac{\sigma_{\mu}^2}{\sigma_{\mu}^2 + \sigma_{\nu}^2} s_{t+1} \\
\hat{\sigma}_{t+1}^2 &= \frac{\sigma_{\nu}^2}{1 + \sigma_{\nu}^2} s_{t+1} \\
z_{t+1} &= \omega_{z} + \phi_{z} z_{t} + \sigma_{z} \bar{e}_{z,t+1}
\end{align*}
\hfill (25)
$$

**Proposition 4.** The decision problem associated with the Bellman equation (13) adjusted by the specifications given in equations (19) to (25) implies the following necessary intertemporal condition for optimal information acquisition $\iota$:

$$E_t \left[ m_{t+1} \Pi_{t+1} + \delta_{t+1} MC_{t+1} \left( \hat{\sigma}_{t+2}^2 \right) \right] = 1,$$

\( \text{The results do not hinge on this specific functional form. More generally, the results hold for any strictly increasing, weakly convex function } n(\iota, h), \text{ with } \partial n(\iota, h) / \partial h |_{h=0} \leq 0 \text{ for } h \geq 0. \) Merely for the purpose of simplifying the final expression, Proposition 4 also uses the special relation $\partial n(\iota, h) / \partial h = -\partial n(\iota, h) / \partial h$ that holds under this quadratic functional specification (see Appendix).
where \( m_{t+1} = \frac{\beta_t F'(c_{t+1})}{F'(c_t)} \) denotes the stochastic discount factor, and

\[
\Pi_{t+1} = \frac{1}{2} \left( \frac{s_{t+1} G(k_{t+1}) - Cov_{t+1}[a_{t+2}, m_{t+2} Y_{t+2}]}{Var_{t+1}[a_{t+2}]} \right)
\]

\[
MC_t (\sigma^2_{t+1}) = -n_t (t_t, h_t) \frac{\partial \delta_{t+1} \left( \sigma^2_{\mu,t}, t_t \right)}{\partial t_t}
\]

\[
\delta_{t+1} = \delta_h \frac{\partial \sigma^2_{t+2} \left( \sigma^2_{\mu,t+1}, t_{t+1} \right)}{\partial t_{t+1}} \frac{\partial \delta_{t+1} \left( \sigma^2_{\mu,t}, t_t \right)}{\partial t_t}
\]

\[
\frac{\partial \delta^2_{t+1} \left( \sigma^2_{\mu,t}, t_t \right)}{\partial t_t} = -\sigma^2_{\mu,t} \left( 1 + \frac{\sigma^2_{\mu,t}}{\sigma^2_{a,t}} \right)^{-2}
\]

Under linear utility with exogenous discount factors \( \beta \), and \( G(k_t) = k^{1-\eta} \) with \( 0 < \eta < 1 \), the following equation holds

\[
E_t \left[ \Pi_{t+1} \right] = E_t \left[ \frac{k_{t+1}}{2\eta} \right]
\]

\[
= E_t \left[ \frac{((1-\eta) \beta E_{t+1}[A_{t+2}])^{1/2}}{2\eta} \right]
\]

\[
= \frac{(1-\eta) \beta}{2\eta} \left\{ \phi + \omega_{a,t+1} + \frac{1}{2} (\sigma^2_{t+1} + \sigma^2_{\mu,t+1}) + \frac{\sigma^2_{\mu,t}}{Var(s_{t+1})} \right\}
\]

Proof. See Appendix.

The structure of the optimality condition in Proposition 4 is similar to the ones presented in the previous Propositions 1 and 3. In particular, for the case of linear utility, the formula provided for \( E_t \left[ \Pi_{t+1} \right] \) is identical to the one given in Proposition 3. The benefit of the setup considered in Proposition 4 is that the results do not hinge on the approximation assumptions made under Propositions 2 and 3. This is due to the fact that persistence in skill, as modeled in this Section, provides a more tractable intertemporal link than persistence in the hidden productivity state.

It is worth noting that given \( \omega_{a,t} \) and \( \sigma^2_{\mu,t} \) are time invariant (i.e., under i.i.d. productivity) there is a constant optimal steady state level for \( t \) that is implicitly determined by the provided optimality condition.

Limitations and Policy Implications

The framework remains agnostic with respect to the relative importance of the various economic forces that determine the parameters of the information acquisition technology. Which economic factors drive the level of marginal cost and which have an impact on the intertemporal scalability of information...
acquisition? Which transition policies would be required to alter information acquisition technologies of a financial system and how costly are these transitions? These questions call for a more explicit analysis of micro-economic information structures and are not addressed by this paper. They are subject of an important literature that explicitly studies different market and intermediation structures.

Yet, the framework may give guidance on the properties of an information acquisition technology that are of relevance. For example, low persistence in the process that governs technology risk exposures of an economy increases the relative value of flexibility enhancing policies, such as reductions of labor market rigidities. In a cost-benefit analysis of different system reform policies the framework might help identifying the relative benefits from altering different channels of a financial system that influence information acquisition.

3 Conclusion

In this paper I examine the link between optimal information acquisition and conditional expectations on future asset market dynamics. The presented model predicts that investors that are Bayesian learners acquire more information in times when future capital productivity is expected to be high, expected returns are low, new technologies are expected to have a persistent impact on productivity, and future investment scalability is expected to be high. Economic mechanisms that decrease the scalability of investment diminish the marginal gains from information acquisition since they restrict the flexibility to respond to news.

References


A Appendix

A.1 Decision Problem

Let $k_t$ be risky investment, $c_t$ consumption, $i_t$ learning-effort, and $F_t$ the decision maker’s information set at time $t$ which includes the history of signals.
denoted by \((a^t, s^t) \equiv \{(a_t, s_t), (a_{t-1}, s_{t-1}), (a_{t-2}, s_{t-2}), \ldots\}\). The production technology is given by
\[ Y_{t+1} = A_{t+1}G_t (k_t) \]
where \(Y_{t+1}\) denotes output at time \(t+1\), \(G_t\) is a strictly increasing, concave, and \(\mathcal{F}_t\)-measurable function, and \(A_{t+1}\) is stochastic conditional on \(\mathcal{F}_t\) and further specified below.

Let \(\beta_t \in (0, 1)\) be a discount factor and \(F (c_{t+j})\) a strictly increasing and concave one-period utility function, and \(y_t \equiv \log Y_t\). Given observed initial values \((y_0, \mu_0, \hat{\sigma}_{p,0}^2, z_0)\) and an initial value of the hidden state \(\mu_0\) that is unobserved by the decision maker, we consider the sequential problem to maximize
\[ F (c_t) + E_t \sum_{j=1}^{\infty} \left( \prod_{i=0}^{j-1} \beta_{t+i} \right) F (c_{t+j}) \]
subject to
\[ c_t = Y_t - k_t - n_t (t_t) \]
\[ y_{t+1} = a_{t+1} + \log G_t (k_t) \]
\[ a_{t+1} = \log A_{t+1} = \omega_{a,t} + \mu_{t+1} + \sigma_{a,t} \varepsilon_{a,t+1} \]
\[ s_{t+1} = \mu_{t+1} + \frac{\sigma_{v}}{\sqrt{t_t}} \varepsilon_{v,t+1} \]
\[ \mu_{t+1} = \omega_{\mu,t} + \phi_{\mu} \mu_t + \sigma_{\mu,t} \varepsilon_{\mu,t+1}, \quad 0 \leq \phi_{\mu} < 1 \]
\[ \hat{\mu}_{t+1} = \omega_{\mu,t} + \phi_{\mu} \frac{\sigma_{\mu,t}^2 \hat{\mu}_t + \hat{\sigma}_{\mu,t}^2 (a_{t+1} - \omega_{a,t}) + \hat{\sigma}_{\varepsilon,t}^2 s_{t+1}}{(\hat{\sigma}_{\mu,t}^2 + \hat{\sigma}_{\varepsilon,t}^2)} \]
\[ \hat{\sigma}_{p,t+1}^2 = \frac{\phi_{\mu}^2 \sigma_{\mu,t}^2}{(\hat{\sigma}_{\mu,t}^2 + \sigma_{\varepsilon,t}^2 + \hat{\sigma}_{\varepsilon,t}^2)} \]
\[ \hat{\sigma}_{\mu,t}^2 = \sigma_{\mu,t}^2 + \sigma_{\varepsilon,t}^2 \]
\[ z_{t+1} = \omega_z + \phi_z z_t + \sigma_z \varepsilon_{z,t+1} \]
\[ \sigma_{\mu,t}^2 = \sigma_{\mu,t}^2 (z_t) > 0 \]
\[ k_t \geq 0 \]
\[ t_t, Y_t > 0 \]

where \(E [\cdot | \mathcal{F}_t]\) denotes the mathematical expectation, given the information set \(\mathcal{F}_t\), \(n_t (t_t)\) denotes a \(\mathcal{F}_t\)-measurable, strictly increasing and convex function which satisfies \(n_t (t_t))_{t_t=0} = 0\) for all \(t\), \(\sigma_{\mu,t}^2 (z_t)\) is a \(\mathcal{F}_t\)-measurable, positive-valued function of the state \(z_t\), and \(\varepsilon_t = (\varepsilon_{a,t}, \varepsilon_{v,t}, \varepsilon_{\mu,t}, \varepsilon_{z,t})'\) an i.i.d. Gaussian shock process with mean zero and identity covariance matrix. The process \(\{\mu_t\}\) is not observed by the decision maker. Associated with this problem is the Bellman equation
\[ V (x_t) = \max_{u_t = (k_t, z_t)'} \{ F (c_t) + \beta_t E [V (x_{t+1}) | \mathcal{F}_t] \} \]
with the state vector \( \mathbf{x}_t \equiv (y_t, \hat{\mu}_t, \hat{\sigma}^2_{p,t}, z_t)' \), the vector of controls \( \mathbf{u}_t \equiv (k_t, \iota_t)' \), and where \( V(\mathbf{x}_t) \) denotes the optimal value of the objective function starting from state \( \mathbf{x}_t \). The state vector transition equations \( \mathbf{x}_{t+1} = \mathbf{h}(\mathbf{x}_t, \mathbf{u}_t, a_{t+1}, s_{t+1}, \varepsilon_{z,t+1}) \) may be written as follows:

\[
\begin{align*}
y_{t+1} & = a_{t+1} + \log G_t(k_t) \\
\hat{\mu}_{t+1} & = \phi_\mu \frac{a_{t+1} - \omega_{a,t+1} s_{t+1} + \frac{\sigma^2}{\sigma^2_t} \hat{\mu}_t}{1 + t + \frac{\sigma^2}{\sigma^2_t}} \\
\hat{\sigma}^2_{p,t+1} & = \phi_\sigma^2 \frac{\sigma^2}{1 + t + \frac{\sigma^2}{\sigma^2_t}} \\
z_{t+1} & = \omega_z + \phi_z z_t + \sigma_z \varepsilon_{z,t+1}
\end{align*}
\]

The first-order necessary conditions for the problem on the right side of equation (27) is

\[
\begin{align*}
\int \frac{\partial \mathbf{x}'_{t+1}}{\partial \mathbf{u}_t} \frac{\partial V(\mathbf{x}_{t+1})}{\partial \mathbf{x}_{t+1}} f_{a,s,\varepsilon z}(a_{t+1}, s_{t+1}, \varepsilon_{z,t+1}|\mathbf{x}_t, \mathbf{u}_t) \, d\lambda(a_{t+1}, s_{t+1}, \varepsilon_{z,t+1}) + \\
\int V(\mathbf{x}_{t+1}) \frac{\partial f_{a,s,\varepsilon z}(a_{t+1}, s_{t+1}, \varepsilon_{z,t+1}|\mathbf{x}_t, \mathbf{u}_t)}{\partial \mathbf{u}_t} \, d\lambda(a_{t+1}, s_{t+1}, \varepsilon_{z,t+1}) \\
= \frac{F'(c_t)}{\beta_t} \left( \frac{1}{n'_t(t_t)} \right)
\end{align*}
\]

with Lebesgue measure \( \lambda \) and \( f_{a,s,\varepsilon z}(a_{t+1}, s_{t+1}, \varepsilon_{z,t+1}|\mathcal{F}_t) \) denoting the joint conditional probability density function of the vector \( (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})' \). Note that due to the given independence assumptions one may write

\[
\frac{\partial f_{a,s,\varepsilon z}(a_{t+1}, s_{t+1}, \varepsilon_{z,t+1}|\mathbf{x}_t, \mathbf{u}_t)}{\partial \mathbf{u}_t} = \begin{pmatrix} 0 & \frac{\partial f_{a,s}(a_{t+1}, s_{t+1}|\mathbf{x}_t, \mathbf{u}_t)}{\partial \mathbf{u}_t} \end{pmatrix} \cdot f_{\varepsilon z}(\varepsilon_{z,t+1}|\mathcal{F}_t).
\]

Further,

\[
\frac{\partial \mathbf{x}'_{t+1}}{\partial \mathbf{u}_t} = \begin{pmatrix} \frac{\partial \mathbf{h}(\mathbf{x}_t, \mathbf{u}_t, a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})'}{\partial \mathbf{u}_t} \\
\end{pmatrix} = \begin{pmatrix} G'_t(k_t) \\
0 \frac{\partial \hat{\mu}_{t+1}}{\partial \mathbf{u}_t} \frac{\partial \hat{\sigma}^2_{p,t+1}}{\partial \mathbf{u}_t} \end{pmatrix}
\]

**Euler Equation**

By the envelope theorem one obtains

\[
\frac{\partial V(\mathbf{x}_t)}{\partial y_t} = F'(c_t) \cdot Y_t
\]

25
which yields the familiar Euler equation

\[
\beta_t E \left[ \frac{G_t'(k_t)}{G_t(k_t)} \frac{\partial V'(x_{t+1})}{\partial y_{t+1}} \mid F_t \right] = F'(c_t)
\]

\[
\beta_t E \left[ A_{t+1}G_t'(k_t) F'(c_{t+1}) \mid F_t \right] = F'(c_t)
\]

A.2 Basic Formulas

This section gives the derivation of some basic formulas that are used in the derivation of the first-order necessary condition for \( \iota_t \).

Kalman Filter

The hidden state \( \mu_{t+1} \) evolves according to the stationary Markov process

\[
\mu_{t+1} = \omega_{\mu,t} + \phi_{\mu} \mu_t + \sigma_{\mu,t} \varepsilon_{\mu,t+1},
\]

the signals are given by

\[
\begin{pmatrix}
    a_{t+1} \\
    s_{t+1}
\end{pmatrix} = \begin{pmatrix}
    \omega_{a,t} \\
    0
\end{pmatrix} + \mu_{t+1} \cdot 1 + \begin{pmatrix}
    \sigma_{v,t} \\
    0
\end{pmatrix} \begin{pmatrix}
    \varepsilon_{a,t+1} \\
    \varepsilon_{s,t+1}
\end{pmatrix},
\]

where \( \mathbf{1} \) denotes a \( 2 \times 1 \) vector of ones, \( (\varepsilon_{a,t}, \varepsilon_{\mu,t}, \varepsilon_{s,t})' \) is an i.i.d. Gaussian shock process with mean zero and identity covariance matrix, and \( \iota_t > 0 \). Define the notation

\[
\hat{\sigma}^2_t \equiv \text{Var} [\mu_{t+1} | F_t]
\]

\[
\hat{\mu}_t \equiv E [\mu_{t+1} | F_t]
\]

where \( F_t \) denotes the information set at time \( t \) which contains the history of signals \( (a^t, s^t) \). Further let \( N (\hat{\mu}_0, \hat{\sigma}^2_0) \) be the initial distribution of \( \mu \). The conditional moments \( (\hat{\mu}_t, \hat{\sigma}^2_t) \) are given by the recursion

\[
\hat{\mu}_t = \omega_{\mu,t-1} + \phi_{\mu} \frac{\sigma_{v,t-1}^2 \hat{\mu}_{t-1} + \hat{\sigma}^2_{t-1} (a_t - \omega_{a,t-1}) + \hat{\sigma}^2_{v,t-1} s_t}{\hat{\sigma}^2_{t-1} + \hat{\sigma}^2_{v,t-1} + \sigma_{v,t-1}^2}
\]

\[
\hat{\sigma}^2_t = \frac{\sigma_{v,t-1}^2}{\sigma_{\mu,t}^2 + \frac{\hat{\sigma}^2_{v,t-1}^2}{\hat{\sigma}^2_{t-1} + \sigma_{v,t-1}^2 + \hat{\sigma}^2_{v,t-1}^2}}.
\]
Some Useful Derivatives

\[
\frac{\partial \mu_{t+1} (a_{t+1}, s_{t+1}, \mu_t, \sigma^2_t, t_t)}{\partial t_t}
= \frac{\partial \left\{ \omega_{\mu, t} + \phi_{\mu} \sigma^2_{\nu, t} \mu_{t+1} + \sigma^2_t (a_{t+1} - \omega_u, t) + \sigma^2_t s_{t+1} \right\}}{\partial t_t}
= \frac{\phi_{\mu} \sigma^2_t (s_{t+1} - \mu_t) - \sigma^2_t (a_{t+1} - \omega_u, t - \mu_t)}{\left( \sigma^2_t + \sigma^2_t s_{t+1} \right)^2}
\]

\[
\frac{\partial^2 \mu_{t+1}}{\partial t_t^2} = -\frac{\phi_{\mu} \sigma^2_{\nu, t}}{\left( \sigma^2_t + \sigma^2_t s_{t+1} \right)^2}
\]

\[
\frac{\partial^2 \mu_{t+1}}{\partial s_{t+1}} = \phi_{\mu} \sigma^2_t (\sigma^2_{\nu, t} + \sigma^2_t)
\]

\[
\frac{\partial \mu_{t+1} (a_{t+1}, s_{t+1}, \mu_t, \sigma^2_t, t_t)}{\partial \mu_t}
= \frac{\partial \left\{ \omega_{\mu, t} + \phi_{\mu} \sigma^2_{\nu, t} \mu_{t+1} + \sigma^2_t (a_{t+1} - \omega_u, t) + \sigma^2_t s_{t+1} \right\}}{\partial \mu_t}
= \frac{\phi_{\mu} \sigma^2_{\nu, t}}{\left( \sigma^2_t + \sigma^2_t s_{t+1} \right)^2}
\]

\[
\frac{\partial \mu_{t+1} (a_{t+1}, s_{t+1}, \mu_t, \sigma^2_t, t_t)}{\partial \sigma^2_t}
= \frac{\partial \left\{ \omega_{\mu, t} + \phi_{\mu} \sigma^2_{\nu, t} \mu_{t+1} + \sigma^2_t (a_{t+1} - \omega_u, t) + \sigma^2_t s_{t+1} \right\}}{\partial \sigma^2_t}
= \frac{\phi_{\mu} \sigma^2_{\nu, t} (a_{t+1} - (\mu_t + \omega_u, t)) + t_t (s_{t+1} - \mu_t)}{\left( \sigma^2_t + \sigma^2_t s_{t+1} + \sigma^2_t \right)^2}
\]

\[
\frac{\partial^2 \sigma^2_{t+1} (\sigma^2_t, t_t)}{\partial \sigma^2_t} = \frac{\partial \left\{ \sigma^2_{\nu, t} + \phi_{\nu} \sigma^2_{\nu, t-1} \right\}}{\partial \sigma^2_t}
= \frac{\phi_{\nu} \sigma^4_t (\sigma^2_t + \sigma^2_t + \sigma^2_t t_t)^2}{\left( \sigma^2_t + \sigma^2_t t_t \right)^2}
\]
\[ \frac{\partial \hat{\sigma}_t^2 (\hat{\sigma}_t^2, \xi_t)}{\partial \xi_t} = \left\{ \sigma_{\mu,t}^2 + \phi^2 \left( \hat{\sigma}_t^2 - 1 + \sigma_{v,t}^2 \right) \right\} \]

\[ = -\phi^2 \frac{\sigma_{v,t}^2}{\left( 1 + \xi_t + \frac{\sigma_{v,t}^2}{\hat{\sigma}_t^2} \right)^2} \]

\[ \frac{\partial^2 \hat{\sigma}_t^2 (\hat{\sigma}_t^2, \xi_t)}{\partial \xi_t^2} = \frac{2 \phi^2 \sigma_{v,t}^2 \left( 1 + \xi_t + \frac{\sigma_{v,t}^2}{\hat{\sigma}_t^2} \right)^{-3}}{1 + \xi_t + \frac{\sigma_{v,t}^2}{\hat{\sigma}_t^2}} \]

\[ \frac{\partial^2 \hat{\sigma}_t^2 (\hat{\sigma}_t^2, \xi_t)}{\partial \xi_t \partial \hat{\sigma}_t^2} = \frac{2}{\left( 1 + \xi_t + \frac{\sigma_{v,t}^2}{\hat{\sigma}_t^2} \right)} \]

A.3 Lemmas

**Lemma 1.** Let \( f (a_{t+1}, s_{t+1} | \mathcal{F}_t) \) denote the conditional joint density of the vector \((a_{t+1}, s_{t+1})'\). Then the following equality holds:

\[ \frac{\partial f (a_{t+1}, s_{t+1} | \mathcal{F}_t)}{\partial \xi_t} \frac{\partial f (a_{t+1}, s_{t+1} | \mathcal{F}_t)}{\partial \hat{\sigma}_t^2} \]

\[ = -\frac{2 \sigma_{v,t}^2 \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 + \xi_t \hat{\sigma}_t^2 \right)^2}{\left( \left( (a_{t+1} - E [a_{t+1} | \mathcal{F}_t]) \hat{\sigma}_t^2 - (s_{t+1} - E [s_{t+1} | \mathcal{F}_t]) (\hat{\sigma}_t^2 + \sigma_{v,t}^2) \right)^2 - Var [a_{t+1} \hat{\sigma}_t^2 - s_{t+1} \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right) | \mathcal{F}_t] \}} \]

**Proof.** Conditional on \( \mathcal{F}_t \) the vector \((a_{t+1}, s_{t+1})'\) is jointly normal with variance-covariance matrix:

\[ \text{Var} (a_{t+1}, s_{t+1} | \mathcal{F}_t) = \begin{pmatrix} \hat{\sigma}_t^2 + \sigma_{v,t}^2 & \hat{\sigma}_t^2 \\ \hat{\sigma}_t^2 & \hat{\sigma}_t^2 + \sigma_{v,t}^2 \end{pmatrix} \]

with

\[ \text{det} [\text{Var} (a_{t+1}, s_{t+1} | \mathcal{F}_t)] = \frac{\sigma_{v,t}^2}{\xi_t} (\hat{\sigma}_t^2 + \sigma_{v,t}^2 + \xi_t \hat{\sigma}_t^2) \]

and

\[ \text{Var} (a_{t+1}, s_{t+1} | \mathcal{F}_t)^{-1} = \frac{\xi_t}{\sigma_{v,t}^2} \begin{pmatrix} \hat{\sigma}_t^2 + \sigma_{v,t}^2 & -\hat{\sigma}_t^2 \\ -\hat{\sigma}_t^2 & \hat{\sigma}_t + \sigma_{v,t}^2 \end{pmatrix} \]

\[ \left( \hat{\sigma}_t^2 + \xi_t \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right) \].
Define

\[
\begin{align*}
\frac{\partial f(a_{t+1}, s_{t+1} | \mathcal{F}_t)}{\partial \xi_{t+1}} = \frac{\partial}{\partial \xi_{t+1}} 2\pi \det \left[ \var{a_{t+1}, s_{t+1} | \mathcal{F}_t} \right]^{\frac{1}{2}} f(a_{t+1}, s_{t+1} | \mathcal{F}_t)
\end{align*}
\]

where the following definition applies:

\[
\begin{align*}
\var{a_{t+1}, s_{t+1} | \mathcal{F}_t} &= \frac{1}{2} \left( a_{t+1} - E \left[ a_{t+1} | \mathcal{F}_t \right] \right)^2 \var{a_{t+1}, s_{t+1} | \mathcal{F}_t}^{-1} \\
&= \frac{2\sigma^2_{v,t} (\hat{\sigma}^2_t + \hat{\sigma}^2_{v,t})}{1 - \left( \mu_t \hat{\sigma}^2_t + \sigma^2_{v,t} \right)} \left( a_{t+1} - E \left[ a_{t+1} | \mathcal{F}_t \right] \right)^2 + 2\sigma^2_{v,t} \left( s_{t+1} - E \left[ s_{t+1} | \mathcal{F}_t \right] \right)^2 - 2\sigma^2_{v,t} \left( a_{t+1} - E \left[ a_{t+1} | \mathcal{F}_t \right] \right) \left( s_{t+1} - E \left[ s_{t+1} | \mathcal{F}_t \right] \right)
\end{align*}
\]

To determine \( \frac{\partial f(a_{t+1}, s_{t+1} | \mathcal{F}_t)}{\partial \xi_{t+1}} \), consider the derivative

\[
\begin{align*}
\frac{\partial \var{a_{t+1}, s_{t+1} | \mathcal{F}_t}}{\partial \xi_{t+1}} &= \frac{1}{2} \left( \hat{\sigma}^2_t (a_{t+1} - E \left[ a_{t+1} | \mathcal{F}_t \right]) - (\hat{\sigma}^2_t + \sigma^2_{v,t}) (s_{t+1} - E \left[ s_{t+1} | \mathcal{F}_t \right]) \right)^2 \\
&= \frac{1}{2} \left( \hat{\sigma}^2_t (a_{t+1} - E \left[ a_{t+1} | \mathcal{F}_t \right]) \right)^2 - \var{a_{t+1}, s_{t+1} | \mathcal{F}_t} \left( \hat{\sigma}^2_t + \sigma^2_{v,t} \right)
\end{align*}
\]

Note that

\[
\begin{align*}
&\frac{1}{2} \left( \hat{\sigma}^2_t + \sigma^2_{v,t} \right) \\
&= \frac{1}{2} \var{a_{t+1} \hat{\sigma}^2_t - s_{t+1} (\hat{\sigma}^2_t + \sigma^2_{v,t}) | \mathcal{F}_t} \\
&= \var{a_{t+1} \hat{\sigma}^2_t - s_{t+1} (\hat{\sigma}^2_t + \sigma^2_{v,t}) | \mathcal{F}_t} \left( \hat{\sigma}^2_t + \sigma^2_{v,t} \right)
\end{align*}
\]

where

\[
\begin{align*}
\var{a_{t+1} \hat{\sigma}^2_t - s_{t+1} (\hat{\sigma}^2_t + \sigma^2_{v,t}) | \mathcal{F}_t} &= \hat{\sigma}^2_t (\hat{\sigma}^2_t + \sigma^2_{v,t}) + (\hat{\sigma}^2_t + \sigma^2_{v,t})^2 \left( \frac{\hat{\sigma}^2_t}{t_t} + \frac{\sigma^2_{v,t}}{t_t} \right) - 2\hat{\sigma}^2_t (\hat{\sigma}^2_t + \sigma^2_{v,t}) \\
&= \sigma^2_{v,t} (\hat{\sigma}^2_t + \sigma^2_{v,t}) \left( \frac{\hat{\sigma}^2_t}{t_t} + \frac{\sigma^2_{v,t}}{t_t} + \hat{\sigma}^2_t \right)
\end{align*}
\]

29
In addition, one may write
\[
\frac{1}{2\pi} \det \left[ \text{Var} \left( a_{t+1}, s_{t+1} | F_t \right) \right]^{0.5} = \frac{1}{2\pi} \left( \frac{\sigma^2_{v,t} (\hat{\sigma}_t^2 + \sigma^2_{v,t} + \dot{\sigma}_t^2 \xi_t)}{\xi_t} \right)^{-0.5}
\]

Taking the derivative w.r.t. \( \xi_t \) yields
\[
\frac{\partial}{\partial \xi_t} \left\{ \frac{1}{2\pi} \det \left[ \text{Var} \left( a_{t+1}, s_{t+1} | F_t \right) \right]^{0.5} \right\} = \frac{\partial}{\partial \xi_t} \left\{ \frac{1}{2\pi} \left( \frac{\sigma^2_{v,t} (\hat{\sigma}_t^2 + \sigma^2_{v,t} + \dot{\sigma}_t^2 \xi_t)}{\xi_t} \right)^{-0.5} \right\}
\]
\[
= \frac{\partial}{\partial \xi_t} \left\{ \frac{\sigma^2_{v,t} (\hat{\sigma}_t^2 + \sigma^2_{v,t} + \dot{\sigma}_t^2 \xi_t)}{2\pi (\hat{\sigma}_t^2 + \sigma^2_{v,t} + \dot{\sigma}_t^2 \xi_t)} \right\}
\]
\[
= \frac{1}{2\pi} \det \left[ \text{Var} \left( a_{t+1}, s_{t+1} | F_t \right) \right]^{0.5}
\]

Taking these results together yields the stated relation.

Lemma 2.
\[
\frac{\partial f(a_{t+1}, s_{t+1} | F_t)}{\partial \hat{\mu}_t} = \frac{(a_{t+1} - \hat{\mu}_t - \omega_{a,t}) + \xi_t (s_{t+1} - \hat{\mu}_t)}{\hat{\sigma}_t^2 + \sigma^2_{v,t} + \dot{\sigma}_t^2 \xi_t}.
\]

Proof. To determine \( \frac{\partial f(a_{t+1}, s_{t+1} | F_t)}{\partial \hat{\mu}_t} \) consider the derivative
\[
\frac{\partial Y (a_{t+1}, s_{t+1}, \xi_t, \hat{\sigma}_t, \hat{\mu}_t)}{\partial \hat{\mu}_t} = \frac{(a_{t+1} - \hat{\mu}_t - \omega_{a,t}) + \xi_t (s_{t+1} - \hat{\mu}_t)}{\hat{\sigma}_t^2 + \sigma^2_{v,t} + \dot{\sigma}_t^2 \xi_t}
\]
The result directly follows.

Lemma 3.
\[
\frac{\partial f(a_{t+1}, s_{t+1} | F_t)}{\partial \sigma^2_{v,t}} = \frac{1}{2} \left( \frac{(a_{t+1} - E [a_{t+1} | F_t]) + \xi_t (s_{t+1} - E [s_{t+1} | F_t])}{(\hat{\sigma}_t^2 + \xi_t \sigma^2_{v,t} + \dot{\sigma}_t^2 \xi_t)^2} - \frac{1}{2} \frac{\text{Var} [a_{t+1} + \xi_t s_{t+1} | F_t]}{(\hat{\sigma}_t^2 + \sigma^2_{v,t} + \dot{\sigma}_t^2 \xi_t)^2} \right)
\]
Thus, one may verify the result

\[
\frac{\partial Y (a_{t+1}, s_{t+1}, \xi_t, \hat{\sigma}_t^2, \hat{\mu}_t)}{\partial \hat{\sigma}_{p,t}} = \frac{1}{2} \left( (a_{t+1} - E [a_{t+1} | F_t]) + \xi_t (s_{t+1} - E [s_{t+1} | F_t]) \right)^2 \left( \hat{\sigma}_t^2 + \xi_t \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right)^2
\]

and

\[
\frac{\partial}{\partial \hat{\sigma}_{p,t}} \left\{ \frac{1}{2 \pi \det [\text{Var} (a_{t+1}, s_{t+1} | F_t)]^{0.5}} \right\}
= \frac{\partial}{\partial \hat{\sigma}_t} \left\{ \frac{1}{2 \pi} \left( \frac{\sigma_{v,t}^2 (\hat{\sigma}_t^2 + \sigma_{v,t}^2 + \hat{\sigma}_t^2 \xi_t)}{\xi_t} \right)^{-\frac{1}{2}} \left( \frac{1}{2} \right) (1 + \xi_t) \right\}
= \frac{1}{2 \pi \det [\text{Var} (a_{t+1}, s_{t+1} | F_t)]^{0.5}} \frac{1}{2} \left( \frac{1}{\sigma_{v,t}^2 + \sigma_{v,t}^2 + \hat{\sigma}_t^2 \xi_t} \right)
= \frac{1}{2 \pi \det [\text{Var} (a_{t+1}, s_{t+1} | F_t)]^{0.5}} \frac{1}{2} \left( \frac{1}{\hat{\sigma}_t^2 + \sigma_{v,t}^2 + \hat{\sigma}_t^2 \xi_t} \right)
\]

Thus, one may verify the result

\[
\frac{\partial f (a_{t+1}, s_{t+1} | F_t)}{\partial \hat{\sigma}_t^2} = \frac{1}{2} \left( (a_{t+1} - E [a_{t+1} | F_t]) + \xi_t (s_{t+1} - E [s_{t+1} | F_t]) \right)^2 \left( \hat{\sigma}_t^2 + \xi_t \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right)^2
\]

where

\[
(1 + \xi_t) \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 + \xi_t \hat{\sigma}_t^2 \right) = \text{Var} (a_{t+1} + \xi_t s_{t+1} | F_t)
\]

\[\square\]

The following Lemmas make use of the following condition and definitions:

**Condition 1.** The function \( \Omega (x) \) satisfies the following conditions

1. \( \Omega (x) \) is a function of the state vector \( x \) only,
2. \( \Omega (h (a, s, \xi)) \) is everywhere differentiable,
3. \( E_t \left[ \left\| \frac{\partial h(a, s, \xi)}{\partial (a, s, \xi)} \right\| \frac{\partial \Omega(x)}{\partial x} \right] < \infty. \)
It may be verified that

\[
E_t \left[ c_{t,t+1} \right] = E_t \left[ \partial \overline{\mu}_t \right] = E_t \left[ c_{t,t+1}^2 \right] = 0
\]

and

\[
\text{Cov}_t \left[ \zeta_{t,t+1}, (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})' \right] = \left( \begin{array}{c}
\frac{\sigma^2_{v,t}}{\phi\sigma_t^2 \sigma^2_{v,t}} \\
0
\end{array} \right)
\]

and

\[
\text{Cov}_t \left[ c_{t,t+1}^2, (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})' \right] = \left( \begin{array}{c}
0 \\
0
\end{array} \right)
\]

Lemma 4. Let \( \Omega \) be a function that satisfies condition \( \Box \) then

\[
E \left[ \frac{\partial \Omega (x_{t+1})}{\partial y_{t+1}} \mid F_t \right] = E \left[ c_{t,t+1} \Omega (x_{t+1}) \mid F_t \right].
\]

Proof. Since \( \zeta_{t,t+1} \) and the vector \( (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})' \) are jointly normal conditional on \( F_t \) one may write by the multivariate version of Stein’s Lemma

\[
E \left[ \frac{(a_{t+1} - s_{t+1} - \omega_{a,t})}{\sigma^2_{v,t}} \Omega (x_{t+1}) \mid F_t \right]
\]

\[
= \left( \begin{array}{c}
-1 \\
0
\end{array} \right) \cdot \frac{\partial x'_{t+1}}{\partial (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})} E \left[ \frac{\partial \Omega (x_{t+1})}{\partial x_{t+1}} \mid F_t \right],
\]

\[
= E \left[ \frac{\partial \Omega (x_{t+1})}{\partial y_{t+1}} \mid F_t \right]
\]

where

\[
\frac{\partial x'_{t+1}}{\partial (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})} = \left( \begin{array}{ccc}
1 & \frac{\phi\sigma^2_t}{\sigma_t^2 + \sigma^2_{x,t} + \sigma^2_{v,t}} & 0 \\
0 & \frac{\phi\sigma^2_{x,t}}{\sigma_t^2 + \sigma^2_{x,t} + \sigma^2_{v,t}} & 0 \\
0 & 0 & \sigma_z
\end{array} \right).
\]
Lemma 5. Let \( \Omega \) be a function that satisfies condition \(|\) then
\[
E \left[ \zeta_{x_{t+1}} \Omega \left( x_{t+1} \right) \bigg| F_t \right] = E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right]
\]

Proof. By the multivariate version of Stein’s Lemma
\[
E \left[ \zeta_{x_{t+1}} \Omega \left( x_{t+1} \right) \bigg| F_t \right] = \text{Cov} \left[ \zeta_{x_{t+1}}, \Omega \left( x_{t+1} \right) \bigg| F_t \right]
\]
\[
= \left( \begin{array}{cc}
\sigma^2_t + \sigma^2_{x,t} + \sigma^2_{z,t} \\
\sigma_t \sigma_{z,t}
\end{array} \right) ' \frac{\partial \Omega'}{\partial (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})} E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right]
\]
\[
= \left( \begin{array}{c}
\sigma_t^2 + \sigma^2_{x,t} \\
0
\end{array} \right) \frac{\partial \Omega'}{\partial (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})} E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right]
\]
\[
= \left( \begin{array}{c}
\sigma_t^2 + \sigma^2_{x,t} \\
0
\end{array} \right) E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right]
\]

Lemma 6. Let \( \Omega \) be a function that satisfies condition \(|\) then
\[
E \left[ \left( a_{t+1} - E \left[ a_{t+1} \big| F_t \right] \right) \Omega \left( x_{t+1} \right) \bigg| F_t \right] = \left( \sigma_t^2 + \sigma^2_{x,t} \right) E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right] + \phi \sigma_t^2 E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right]
\]

Proof. By the multivariate version of Stein’s Lemma one may write
\[
E \left[ \left( a_{t+1} - E \left[ a_{t+1} \big| F_t \right] \right) \Omega \left( x_{t+1} \right) \bigg| F_t \right] = \left( \sigma_t^2 + \sigma^2_{x,t} \right) \frac{\partial \Omega'}{\partial \mu_{t+1}} E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right]
\]
\[
= \left( \sigma_t^2 + \sigma^2_{x,t} \right) E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right] + \phi \sigma_t^2 E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right]
\]

Lemma 7. Let \( \Omega \) be a function that satisfies condition \(|\) then
\[
E \left[ \left( s_{t+1} - E \left[ s_{t+1} \big| F_t \right] \right) \Omega \left( x_{t+1} \right) \bigg| F_t \right] = \sigma_t^2 E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right] + \phi \sigma_t^2 E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right]
\]

Proof. By the multivariate version of Stein’s Lemma one may write
\[
E \left[ \left( s_{t+1} - E \left[ s_{t+1} \big| F_t \right] \right) \Omega \left( x_{t+1} \right) \bigg| F_t \right] = \sigma_t^2 \frac{\partial \Omega'}{\partial \mu_{t+1}} E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right]
\]
\[
= \sigma_t^2 E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right] + \phi \sigma_t^2 E \left[ \frac{\partial \Omega \left( x_{t+1} \right)}{\partial \mu_{t+1}} \bigg| F_t \right]
\]
Lemma 8. Let $\Omega$ be a function that satisfies condition \([1]\) then

$$E_t \left[ \frac{\partial \mu_t^2}{\partial \mu_{t+1}^2} \right] = E_t \left[ \frac{\partial^2 \Omega (x_{t+1})}{\partial \mu_{t+1}^2} \right]$$

(29)

Proof. Since $\frac{\partial \mu_t^2}{\partial \mu_{t+1}^2}$ is not conditionally jointly normal with the vector $(a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})'$ consider the decomposition

$$((a_{t+1} - E_t [a_{t+1}]) \sigma_t^2 - (s_{t+1} - E_t [s_{t+1}]) (\sigma_t^2 + \sigma_{v,t}^2))^2$$

$$= \left(\begin{array}{c}
(a_{t+1} - E_t [a_{t+1}]) \\
(s_{t+1} - E_t [s_{t+1}])
\end{array}\right)'$$

$$\left(\begin{array}{c}
(\sigma_t^2)^2 (a_{t+1} - E_t [a_{t+1}]) - \sigma_t^2 (\sigma_t^2 + \sigma_{v,t}^2) (s_{t+1} - E_t [s_{t+1}]) \\
(\sigma_t^2 + \sigma_{v,t}^2)^2 (s_{t+1} - E_t [s_{t+1}]) - \sigma_t^2 (\sigma_t^2 + \sigma_{v,t}^2) (a_{t+1} - E_t [a_{t+1}])
\end{array}\right)$$

Then, one may write

$$E_t \left[ V (x_{t+1}) ((a_{t+1} - E_t [a_{t+1}]) \sigma_t^2 - (s_{t+1} - E_t [s_{t+1}]) (\sigma_t^2 + \sigma_{v,t}^2))^2 \right]$$

$$= \text{Cov}_t \left[(\sigma_t^2)^2 a_{t+1} - \sigma_t^2 (\sigma_t^2 + \sigma_{v,t}^2) s_{t+1}, \Omega (x_{t+1}) (a_{t+1} - E_t [a_{t+1}]) \right] +$$

$$\text{Cov}_t \left[(\sigma_t^2 + \sigma_{v,t}^2)^2 s_{t+1} - \sigma_t^2 (\sigma_t^2 + \sigma_{v,t}^2) a_{t+1}, \Omega (x_{t+1}) (s_{t+1} - E_t [s_{t+1}]) \right]$$

First, consider

$$\text{Cov}_t \left[(\sigma_t^2)^2 a_{t+1} - \sigma_t^2 (\sigma_t^2 + \sigma_{v,t}^2) s_{t+1}, \Omega (x_{t+1}) (a_{t+1} - E_t [a_{t+1}]) \right]$$

$$= \text{Cov}_t \left[(\sigma_t^2)^2 a_{t+1} - \sigma_t^2 (\sigma_t^2 + \sigma_{v,t}^2) s_{t+1}, (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})' \right]'$$

$$E_t \left[ \frac{\partial x_{t+1}'}{\partial (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})} \frac{\partial \Omega (x_{t+1})}{\partial x_{t+1}} (a_{t+1} - E_t [a_{t+1}]) \right] +$$

$$\text{Cov}_t \left[(\sigma_t^2)^2 a_{t+1} - \sigma_t^2 (\sigma_t^2 + \sigma_{v,t}^2) s_{t+1}, (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})' \right]'$$

$$E_t \left[ \frac{\partial (a_{t+1} - E_t [a_{t+1}])}{\partial (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})}' \Omega (x_{t+1}) \right]$$

Substituting

$$\text{Cov}_t \left[(\sigma_t^2)^2 a_{t+1} - \sigma_t^2 (\sigma_t^2 + \sigma_{v,t}^2) s_{t+1}, (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})' \right]'$$

$$= \left( - \frac{\sigma_t^2 \sigma_{v,t}^2 (\sigma_t^2 + \sigma_{v,t}^2 + \sigma_{v,t}^2)}{0} \right)'$$

and rearranging yields

$$\text{Cov}_t \left[(\sigma_t^2)^2 a_{t+1} - \sigma_t^2 (\sigma_t^2 + \sigma_{v,t}^2) s_{t+1}, \Omega (x_{t+1}) (a_{t+1} - E_t [a_{t+1}]) \right]$$

$$= - \phi_\mu \sigma_{v,t}^2 \sigma_t^4 E_t \left[ \frac{\partial \Omega (x_{t+1})}{\partial \mu_{t+1}^2} (a_{t+1} - E_t [a_{t+1}]) \right]$$

34
Given that \( \frac{\partial \Omega (x_{t+1})}{\partial \mu_{t+1}} \) satisfies condition [1], one may write by Lemma [6]

\[
E_t \left[ \frac{\partial \Omega (x_{t+1})}{\partial \mu_{t+1}} (a_{t+1} - E_t [a_{t+1}]) \right] 
= (\hat{\sigma}_t^2 + \sigma_{v,t}^2) E_t \left[ \frac{\partial^2 \Omega (x_{t+1})}{\partial y_{t+1} \partial \mu_{t+1}} \right] + \phi_{\mu} \hat{\sigma}_t^2 E_t \left[ \frac{\partial^2 \Omega (x_{t+1})}{\partial \mu_{t+1}^2} \right]
\]

which yields

\[
\text{Cov}_t \left[ \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right)^2 \hat{\sigma}_t^2 \right] (a_{t+1} - E_t [a_{t+1}]) \Omega (x_{t+1}) (s_{t+1} - E_t [s_{t+1}]) 
= -\phi_{\mu} \sigma_{v,t}^2 \hat{\sigma}_t^2 \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right) E_t \left[ \frac{\partial^2 \Omega (x_{t+1})}{\partial y_{t+1} \partial \mu_{t+1}} \right] - \phi_{\mu} \sigma_{v,t}^2 \hat{\sigma}_t^2 E_t \left[ \frac{\partial^2 \Omega (x_{t+1})}{\partial \mu_{t+1}^2} \right]
\]

Next, consider the term

\[
\text{Cov}_t \left[ \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right)^2 s_{t+1} - \hat{\sigma}_t^2 \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right) a_{t+1} \Omega (x_{t+1}) (s_{t+1} - E_t [s_{t+1}]) \right]
= \text{Cov}_t \left[ \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right)^2 s_{t+1} - \hat{\sigma}_t^2 \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right) a_{t+1} \right] (s_{t+1} - E_t [s_{t+1}]) \Omega (x_{t+1})
\]

Substituting

\[
\text{Cov}_t \left[ \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right)^2 s_{t+1} - \hat{\sigma}_t^2 \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right) a_{t+1}, (s_{t+1} - E_t [s_{t+1}]) \right]
= \left( \begin{array}{c} 0 \\ \sigma_{v,t}^2 \left( \sigma_{v,t}^2 + \sigma_{v,t}^2 \right) \\ \hat{\sigma}_t^2 \left( \sigma_{v,t}^2 + \sigma_{v,t}^2 + \lambda \right) \\ 0 \end{array} \right)
\]

and rearranging yields

\[
\text{Cov}_t \left[ \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right)^2 s_{t+1} - \hat{\sigma}_t^2 \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right) a_{t+1}, \Omega (x_{t+1}) (s_{t+1} - E_t [s_{t+1}]) \right]
= \frac{\sigma_{v,t}^2 \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right)}{\sigma_{v,t}^2 + \hat{\sigma}_t^2 + \lambda} \left( \frac{\phi_{\mu} \hat{\sigma}_t^2 \lambda}{\sigma_{v,t}^2 + \hat{\sigma}_t^2 + \lambda} E_t \left[ \frac{\partial \Omega (x_{t+1})}{\partial \mu_{t+1}} (s_{t+1} - E_t [s_{t+1}]) \right] + E_t \left[ \Omega (x_{t+1}) \right] \right)
\]

35
Given that $\frac{\partial \Omega(x_{t+1})}{\partial \mu_{t+1}}$ satisfies condition (1) one may write by Lemma (7)

\[
E_t \left[ \frac{\partial \Omega(x_{t+1})}{\partial \mu_{t+1}} (s_{t+1} - E_t [s_{t+1}]) \right] = \sigma_{t}^2 E_t \left[ \frac{\partial^2 \Omega(x_{t+1})}{\partial y_{t+1} \partial \mu_{t+1}} + \phi_{t} \sigma_{t}^2 E_t \left[ \frac{\partial^2 \Omega(x_{t+1})}{\partial \mu_{t+1}^2} \right] \right]
\]

which yields

\[
Cov_t \left[ (\partial_t^2 + \sigma_{v,t}^2)^2 s_{t+1} - \partial_t^2 (\partial_t^2 + \sigma_{v,t}^2) \right] a_{t+1}, \Omega(x_{t+1}) (s_{t+1} - E_t [s_{t+1}])
\]

\[
= \phi_{t} \sigma_{v,t}^2 \left( \partial_t^2 + \sigma_{v,t}^2 \right) \tilde{\sigma}_t^2 E_t \left[ \frac{\partial^2 \Omega(x_{t+1})}{\partial y_{t+1} \partial \mu_{t+1}} + \sigma_{v,t}^2 \left( \partial_t^2 + \sigma_{v,t}^2 + \nu \partial_t^2 \right) \right] E_t [\Omega(x_{t+1})]
\]

Substituting these results yields

\[
E_t \left[ \Omega(x_{t+1}) \left( (a_{t+1} - E_t [a_{t+1}]) \partial_t^2 - (s_{t+1} - E_t [s_{t+1}]) (\partial_t^2 + \sigma_{v,t}^2) \right)^2 \right]
\]

\[
= \phi_{t} \sigma_{v,t}^2 \left( \partial_t^2 + \sigma_{v,t}^2 \right) \tilde{\sigma}_t^2 E_t \left[ \frac{\partial^2 \Omega(x_{t+1})}{\partial \mu_{t+1}^2} \right] + \frac{Var_t [a_{t+1} \partial_t^2 - s_{t+1} (\partial_t^2 + \sigma_{v,t}^2)] \cdot E_t [\Omega(x_{t+1})]}{\nu}
\]

since

\[
Var_t [a_{t+1} \partial_t^2 - s_{t+1} (\partial_t^2 + \sigma_{v,t}^2)] = \frac{\sigma_{v,t}^2 \cdot (\partial_t^2 + \sigma_{v,t}^2) \cdot (\nu \partial_t^2 + \partial_t^2 + \sigma_{v,t}^2)}{\nu}
\]

Thus, one obtains

\[
E_t \left[ \frac{\partial^2 \Omega(x_{t+1})}{\partial y_{t+1} \partial \mu_{t+1}} \Omega(x_{t+1}) \right] = E_t \left[ \left( (a_{t+1} - E_t [a_{t+1}]) \partial_t^2 - (s_{t+1} - E_t [s_{t+1}]) (\partial_t^2 + \sigma_{v,t}^2) \right)^2 \Omega(x_{t+1}) \right] - \frac{\phi_{t} \sigma_{v,t}^2 \tilde{\sigma}_t^2}{\nu \partial_t^2 + \partial_t^2 + \sigma_{v,t}^2} \cdot \frac{Var_t [a_{t+1} \partial_t^2 - s_{t+1} (\partial_t^2 + \sigma_{v,t}^2)]}{\nu \partial_t^2 + \partial_t^2 + \sigma_{v,t}^2} \cdot E_t [\Omega(x_{t+1})]
\]

\[
= E_t \left[ \frac{\partial^2 \Omega(x_{t+1})}{\partial \mu_{t+1}^2} \right]
\]
Lemma 9. Given the value function $V(x_{t+1})$ satisfies condition 1 the following equality holds:

$$
E \left[ V(x_{t+1}) \frac{\partial f_{a_{t+1}, s_{t+1}}}{\partial \mu_t} \bigg| F_t \right] = E \left[ \frac{\partial V(x_{t+1})}{\partial y_{t+1}} \bigg| F_t \right] + \left( 1 + \nu_t \right) \hat{\sigma}_t^2 \phi_{\mu} E \left[ \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} \bigg| F_t \right]
$$

Proof. Using the previously derived representation of $\frac{\partial f_{a_{t+1}, s_{t+1}}}{\partial \mu_t} F_t$ one may write

$$
E \left[ \frac{\partial f_{a_{t+1}, s_{t+1}}}{\partial \mu_t} \bigg| F_t \right] = E_t \left[ V(x_{t+1}) (a_{t+1} - E_t [a_{t+1}]) \right] + \nu_t E_t [V(x_{t+1}) (s_{t+1} - \hat{\mu}_t)]
$$

Given that $V(x_{t+1})$ satisfies condition 1 one may write by Lemma 6 and Lemma 7:

$$
E \left[ \frac{\partial f_{a_{t+1}, s_{t+1}}}{\partial \mu_t} \bigg| F_t \right] = \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right) E \left[ \frac{\partial V(x_{t+1})}{\partial y_{t+1}} \bigg| F_t \right] + \phi_{\mu} \hat{\sigma}_t^2 E \left[ \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} \bigg| F_t \right]
$$

Substituting these results and gathering common terms yields

$$
E \left[ V(x_{t+1}) \frac{\partial f_{a_{t+1}, s_{t+1}}}{\partial \mu_t} \bigg| F_t \right] = E \left[ \frac{\partial V(x_{t+1})}{\partial y_{t+1}} \bigg| F_t \right] + \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 \right) \phi_{\mu} E \left[ \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} \bigg| F_t \right]
$$

Lemma 10. The shadow value of $\hat{\mu}$ has the following recursive representation:

$$
\frac{\partial V(x_t)}{\partial \hat{\mu}_t} = \beta_t E \left[ \frac{\partial V(x_{t+1})}{\partial y_{t+1}} \bigg| F_t \right] + \phi_{\mu} \beta_t E \left[ \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} \bigg| F_t \right]
$$
Proof. By the envelope theorem one may write \( \frac{\partial V(x_t)}{\partial \mu_t} \) as follows

\[
\begin{align*}
\frac{\partial V(x_t)}{\partial \mu_t} &= \beta_t E \left[ \frac{\partial h'_{t+1}}{\partial \mu_t} \frac{\partial V(x_{t+1})}{\partial x_{t+1}} + V(x_{t+1}) \frac{\partial f(a_{t+1}, s_{t+1} \mid F_t)}{\partial \mu_t} \right]_{\mathcal{F}_t} \\
&= \beta_t E \left[ \frac{\partial \mu_{t+1}}{\partial \mu_t} \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} + V(x_{t+1}) \frac{\partial f(a_{t+1}, s_{t+1} \mid F_t)}{\partial \mu_t} \right]_{\mathcal{F}_t}
\end{align*}
\]

since

\[
\frac{\partial h'_{t+1}}{\partial \mu_t} = \left( \begin{array}{c} 0 \\ \frac{\partial h'_{t+1}}{\partial \mu_t} \\ 0 \\ 0 \end{array} \right),
\]

where

\[
\frac{\partial f(a_{t+1}, s_{t+1} \mid x_t, u_t)}{\partial \mu_t} = \left( \begin{array}{c} (a_{t+1} - \mu_t - \omega_a, t) + \iota_t (s_{t+1} - \mu_t) \\ \sigma^2_t + \sigma^2_{v,t} + \iota_t \sigma^2_{e,t} \end{array} \right)
\]

as defined and derived above. Since, as shown above, \( \frac{\partial \mu_{t+1}}{\partial \mu_t} \) is \( \mathcal{F}_t \)-measurable one may further write

\[
\begin{align*}
\frac{\partial V(x_t)}{\partial \mu_t} &= \beta_t \frac{\partial \mu_{t+1}}{\partial \mu_t} E \left[ \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} \right]_{\mathcal{F}_t} + \beta_t E \left[ \frac{\partial f(a_{t+1}, s_{t+1} \mid F_t)}{\partial \mu_t} \right]_{\mathcal{F}_t} \\
&= \beta_t E \left[ \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} \right]_{\mathcal{F}_t} + \beta_t E \left[ \frac{\partial f(a_{t+1}, s_{t+1} \mid F_t)}{\partial \mu_t} \right]_{\mathcal{F}_t}
\end{align*}
\]

Further, using Lemma (9) and the previously derived representation for \( \frac{\partial \mu_{t+1}}{\partial \mu_t} \),

\[
\frac{\partial \mu_{t+1}}{\partial \mu_t} = \phi\mu \frac{\sigma^2_{v,t}}{\sigma^2_t + \sigma^2_{e,t} + \sigma^2_{v,t}},
\]

one may write equation (30) as follows

\[
\begin{align*}
\frac{\partial V(x_t)}{\partial \mu_t} &= \beta_t \phi \mu \frac{\sigma^2_{e,t}}{\sigma^2_t + \sigma^2_{e,t} + \sigma^2_{v,t}} E \left[ \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} \right]_{\mathcal{F}_t} + \\
&\quad + \beta_t E \left[ \frac{\partial V(x_{t+1})}{\partial y_{t+1}} \right]_{\mathcal{F}_t} + \\
&\quad + \beta_t \phi \mu \frac{\sigma^2_{e,t} + \sigma^2_{v,t}}{\sigma^2_t + \sigma^2_{e,t} + \sigma^2_{v,t}} E \left[ \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} \right]_{\mathcal{F}_t}
\end{align*}
\]

This yields

\[
\begin{align*}
\frac{\partial V(x_t)}{\partial \mu_t} &= \beta_t E \left[ \frac{\partial V(x_{t+1})}{\partial y_{t+1}} \right]_{\mathcal{F}_t} + \phi\mu \beta_t E \left[ \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} \right]_{\mathcal{F}_t}. \quad (31)
\end{align*}
\]
Lemma 11.

\[ E \left[ \frac{\partial \hat{e}_{t+1}^2}{\partial \hat{u}_t} \frac{\partial V \left( x_{t+1} \right)}{\partial \hat{e}_{t+1}} \right] = - \frac{\partial \hat{e}_{t+1}^2}{\partial \hat{u}_t} E \left[ \frac{\partial V \left( x_{t+1} \right)}{\partial \hat{e}_{t+1}} \right] \]  \hspace{0.5cm} (32)

Proof. Note that given the previous definitions one may write

\[ \frac{\partial \hat{e}_{t+1}}{\partial \hat{u}_t} = \frac{\partial \hat{e}_{t+1}}{\partial \hat{u}_t} \cdot \hat{e}_{t+1}. \]

Given that \( \frac{\partial V \left( x_{t+1} \right)}{\partial \hat{e}_{t+1}} \) satisfies condition [1], one directly obtains the stated result by applying Lemma [5]. \( \Box \)

Lemma 12. Define \( V^{RS} \left( x_t, u \left( x_t \right) \right) \) as the right side of equation [27] evaluated at the optimal policy function \( u \left( x_t \right) \). \( \frac{\partial^2 V \left( x_{t+1} \right)}{\partial \hat{u}_t^2} \) has the following representation:

\[ \frac{\partial^2 V \left( x_{t+1} \right)}{\partial \hat{u}_t^2} = \beta_{t+1} E \left[ \frac{\partial^2 V \left( x_{t+2} \right)}{\partial y_{t+2}^2} \right] + \phi_\mu^2 \beta_{t+1} E \left[ \frac{\partial^2 V \left( x_{t+2} \right)}{\partial \hat{e}_{t+2}^2} \right] + 2 \phi_\mu \beta_{t+1} E \left[ \frac{\partial^2 V^{RS} \left( x_{t+1}, u \left( x_{t+1} \right) \right)}{\partial \hat{e}_{t+1} \partial \hat{u}_t} \right] + 2 \phi_\mu \beta_{t+1} E \left[ \frac{\partial^2 V^{RS} \left( x_{t+1}, u \left( x_{t+1} \right) \right)}{\partial \hat{u}_t \partial \hat{u}_t} \right] \left( \frac{\partial u \left( x_{t+1} \right)}{\partial \hat{u}_t} \right)^2. \]  \hspace{0.5cm} (33)

Proof. Moving equation [31] one period forward and differentiating w.r.t. \( \hat{u}_{t+1} \) yields

\[ \frac{\partial^2 V \left( x_{t+1} \right)}{\partial \hat{u}_t^2} = \frac{\partial^2 V^{RS} \left( x_{t+1}, u \left( x_{t+1} \right) \right)}{\partial \hat{u}_t^2} + 2 \frac{\partial^2 V^{RS} \left( x_{t+1}, u \left( x_{t+1} \right) \right)}{\partial \hat{u}_t \partial \hat{u}_t} \frac{\partial u \left( x_{t+1} \right)}{\partial \hat{u}_t} + \frac{\partial^2 V^{RS} \left( x_{t+1}, u \left( x_{t+1} \right) \right)}{\partial \hat{u}_t^2} \left( \frac{\partial u \left( x_{t+1} \right)}{\partial \hat{u}_t} \right)^2. \]

By differentiating [31] w.r.t. \( \hat{u}_{t+1} \) one obtains

\[ \frac{\partial^2 V^{RS} \left( x_{t+1}, u \left( x_{t+1} \right) \right)}{\partial \hat{u}_t^2} = \beta_{t+1} \hat{u}_{t+1} E \left[ \frac{\partial V \left( x_{t+2} \right)}{\partial \hat{e}_{t+2}} \right] + \phi_\mu \beta_{t+1} \frac{\partial E \left[ \frac{\partial V \left( x_{t+2} \right)}{\partial \hat{e}_{t+2}} \right]}{\partial \hat{u}_t^2}. \]  \hspace{0.5cm} (34)

First, consider the term

\[ \frac{\partial E \left[ \frac{\partial V \left( x_{t+2} \right)}{\partial \hat{e}_{t+2}} \right]}{\partial \hat{u}_t^2} \]  \hspace{0.5cm} (35)

\[ = E \left[ \frac{\partial V \left( x_{t+2} \right)}{\partial \hat{e}_{t+2}} \right] \frac{\partial E \left[ \frac{\partial V \left( x_{t+2} \right)}{\partial \hat{e}_{t+2}} \right]}{\partial \hat{u}_t^2} + E \left[ \frac{\partial V \left( x_{t+2} \right)}{\partial \hat{e}_{t+2}} \right] \frac{\partial E \left[ \frac{\partial V \left( x_{t+2} \right)}{\partial \hat{e}_{t+2}} \right]}{\partial \hat{u}_t^2} \right] \left( \frac{\partial u \left( x_{t+1} \right)}{\partial \hat{u}_t} \right)^2. \]
Note that

\[
E \left[ \frac{\partial V(x_{t+2})}{\partial \mu_{t+1}} \right] = E \left[ \frac{\partial h_{t+2}}{\partial \mu_{t+1}} \right] = \frac{\phi_\mu \sigma_{v,t+1}^2}{\bar{\sigma}_{t+1}^2 + \sigma_{v,t+1}^2 + \bar{\sigma}_{t+1}^2} \cdot E \left[ \frac{\partial^2 V(x_{t+2})}{\partial y_{t+2} \partial \mu_{t+2}} \right] \]

since

\[
\frac{\partial h_{t+2}}{\partial \mu_{t+1}} = \left( \frac{\phi_\mu \sigma_{v,t+1}^2}{\bar{\sigma}_{t+1}^2 + \sigma_{v,t+1}^2 + \bar{\sigma}_{t+1}^2}, 0 \right).
\]  

(36)

Next, consider the second addend

\[
E \left[ \frac{\partial f(a_{t+2}, s_{t+2} | x_{t+1}, u_{t+1})}{\partial y_{t+2}} \right] \left[ \frac{\partial V(x_{t+2})}{\partial y_{t+2}} \right] \left| F_{t+1} \right.
\]

\[
= E_t \left[ \frac{\partial f(a_{t+2}, s_{t+2} | x_{t+1}, u_{t+1})}{\partial y_{t+2}} (a_{t+1} - E_t [a_{t+1}] \right) + t_t \cdot E_t \left[ \frac{\partial V(x_{t+2})}{\partial y_{t+2}} (s_{t+1} - \hat{\mu}_t) \right]
\]

Given that \( \frac{\partial V(x_{t+2})}{\partial y_{t+2}} \) satisfies condition (1), one may apply Lemma (6) and Lemma (7) to obtain:

\[
E \left[ (a_{t+2} - E [a_{t+2} | F_{t+1}]) \frac{\partial V(x_{t+2})}{\partial y_{t+2}} \right] \left| F_{t+1} \right.
\]

\[
= (\hat{\sigma}_{t+1}^2 + \sigma_{v,t+1}^2) E_t \left[ \frac{\partial V(x_{t+2})}{\partial y_{t+2}} \right] + \phi_\mu \hat{\sigma}_{t+1}^2 E_t \left[ \frac{\partial^2 V(x_{t+2})}{\partial y_{t+2} \partial y_{t+2}} \right] \left| F_{t+1} \right.
\]

\[
E \left[ (s_{t+2} - E [s_{t+2} | F_{t+1}]) V(x_{t+2}) \right] \left| F_{t+1} \right.
\]

\[
= \hat{\sigma}_{t+1}^2 E_t \left[ \frac{\partial V(x_{t+2})}{\partial y_{t+2}} \right] + \phi_\mu \hat{\sigma}_{t+1}^2 E_t \left[ \frac{\partial^2 V(x_{t+2})}{\partial y_{t+2} \partial y_{t+2}} \right] \left| F_{t+1} \right.
\]

Substituting these results and gathering common terms yields

\[
E \left[ \frac{\partial f(a_{t+2}, s_{t+2} | x_{t+1}, u_{t+1})}{\partial y_{t+2}} (a_{t+1} - E_t [a_{t+1}] \right) + t_t \cdot E_t \left[ \frac{\partial V(x_{t+2})}{\partial y_{t+2}} (s_{t+1} - \hat{\mu}_t) \right]
\]

\[
= \hat{\sigma}_{t+1}^2 E_t \left[ \frac{\partial V(x_{t+2})}{\partial y_{t+2}} \right] + \phi_\mu \hat{\sigma}_{t+1}^2 E_t \left[ \frac{\partial^2 V(x_{t+2})}{\partial y_{t+2} \partial y_{t+2}} \right] \left| F_{t+1} \right.
\]

(37)
Substituting these results into eqn. (35) and gathering common terms yields

$$
\frac{\partial E}{\partial \hat{\mu}_{t+1}} \left[ \frac{\partial V(x_{t+2})}{\partial y_{t+2}} \right] F_{t+1} = E \left[ \frac{\partial^2 V(x_{t+2})}{\partial y_{t+2}^2} \right] F_{t+1} + \phi \mu E \left[ \frac{\partial^2 V(x_{t+2})}{\partial y_{t+2} \partial \hat{\mu}_{t+1}} \right] F_{t+1}
$$

(38)

Next, consider the term

$$
\frac{\partial E}{\partial \hat{\mu}_{t+1}} \left[ \frac{\partial V(x_{t+2})}{\partial \hat{\mu}_{t+1}} \right] F_{t+1}
$$

$$
= E \left[ \frac{\partial V(x_{t+2})}{\partial \hat{\mu}_{t+1}} \right] F_{t+1} + E \left[ \frac{\partial V(x_{t+2})}{\partial \hat{\mu}_{t+1}} \frac{\partial f(a_{t+2}, s_{t+2}|F_{t+1})}{\partial \hat{\mu}_{t+1}} \right] F_{t+1}
$$

where, using (36), the first addend may be written as follows

$$
E \left[ \frac{\partial V(x_{t+2})}{\partial \hat{\mu}_{t+1}} \right] F_{t+1} = E \left[ \frac{\partial h''(t+1)}{\partial \hat{\mu}_{t+1}} \right] F_{t+1}
$$

$$
= E \left[ \frac{\partial V(x_{t+2})}{\partial \hat{\mu}_{t+1}} \frac{\partial f(a_{t+2}, s_{t+2}|F_{t+1})}{\partial \hat{\mu}_{t+1}} \right] F_{t+1}
$$

and Lemma (7) as follows:

$$
E \left[ \frac{\partial V(x_{t+2})}{\partial \hat{\mu}_{t+1}} \right] F_{t+1} = E \left[ \frac{\partial V(x_{t+2})}{\partial \hat{\mu}_{t+1}} \frac{\partial f(a_{t+2}, s_{t+2}|F_{t+1})}{\partial \hat{\mu}_{t+1}} \right] F_{t+1}
$$

Further, analogous to (37), one may write the second addend by Lemma (6) and Lemma (7) as follows:

$$
E \left[ \frac{\partial V(x_{t+2})}{\partial y_{t+2}} \right] F_{t+1} + \phi \mu \frac{\partial^2 V(x_{t+2})}{\partial y_{t+2} \partial \hat{\mu}_{t+1}} F_{t+1}
$$

Substituting the results for the two addends yields

$$
\frac{dE}{\partial \hat{\mu}_{t+1}} \left[ \frac{\partial V(x_{t+2})}{\partial y_{t+2}} \right] F_{t+1} = E \left[ \frac{\partial V(x_{t+2})}{\partial y_{t+2}} \right] F_{t+1} + \phi \mu \left[ \frac{\partial^2 V(x_{t+2})}{\partial y_{t+2}^2} \right] F_{t+1}
$$

(39)
Substituting (39) and (38) into (34) and gathering common terms yields

\[
\frac{\partial^2 V^{RS}}{\partial \mu_{t+1}^2} (x_{t+1}, u(x_{t+1})) = \beta_{t+1} E \left[ \frac{\partial^2 V (x_{t+2})}{\partial y_{t+2}^2} \right] F_{t+1} + \\
\phi^2_{\mu} \beta_{t+1} E \left[ \frac{\partial^2 V (x_{t+2})}{\partial \mu_{t+1}^2} \right] F_{t+1} + \\
2\phi_{\mu} \beta_{t+1} E \left[ \frac{\partial^2 V (x_{t+2})}{\partial y_{t+2} \partial \mu_{t+1}} \right] F_{t+1}
\]

which yields the stated relation.

\[\square\]

**Lemma 13.**

\[
E \left[ V (x_{t+1}) \frac{\partial f(a_{t+1}, s_{t+1} \mid x_t, u_t)}{\partial \sigma^2_{\mu,t}} \right] F_t = 1 \cdot E \left[ \frac{\partial^2 V (x_{t+1})}{\partial y_{t+1}^2} \right] F_t + \\
\frac{1}{2} \phi^2_{\mu} (1 + \iota_t)^2 \sigma^4_{\epsilon,t} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}^2} \right] F_t + \\
\phi_{\mu} (1 + \iota_t) \sigma^2_{\epsilon,t} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1} \partial y_{t+1}} \right] F_t
\]

**Proof.** Note that one may write \( \frac{\partial f(a_{t+1}, s_{t+1} \mid x_t, u_t)}{\partial \sigma^2_{\mu,t}} \) as follows:

\[
\frac{\partial f(a_{t+1}, s_{t+1} \mid x_t, u_t)}{\partial \sigma^2_{\mu,t}} \bigg|_{\sigma^2_{\mu,t} = \sigma^2_{\mu,t} + \Delta} = 1 \cdot \frac{1}{2} \left( \frac{1}{(1 + \iota_t)^2 + \sigma^2_{\epsilon,t}^2} \right)^2 \left\{ \left( a_{t+1} - E [a_{t+1} \mid F_t] \right) + \iota_t (s_{t+1} - E [s_{t+1} \mid F_t]) \right\}
\]

This decomposition is helpful since \( \frac{\partial f(a_{t+1}, s_{t+1} \mid x_t, u_t)}{\partial \sigma^2_{\mu,t}} \) is not conditionally normal. Further, define the vector

\[
\left( \begin{array}{c}
V_{t,1} (x_{t+1}, a_{t+1}) \\
V_{t,2} (x_{t+1}, s_{t+1})
\end{array} \right) = V (x_{t+1}) \left( \begin{array}{c}
(a_{t+1} - E [a_{t+1} \mid F_t]) \\
(s_{t+1} - E [s_{t+1} \mid F_t])
\end{array} \right).
\]
Then one may write

\[
E \left[ V(\mathbf{x}_{t+1}) \right. \frac{\partial f(a_{t+1}, s_{t+1} | \mathbf{x}_t, \mathbf{u}_t)}{\partial \sigma^2_{p,t}} \bigg| \mathcal{F}_t \right] = \frac{1}{2} \left( \frac{1}{\hat{\sigma}^2_t (1 + \theta_t) + \sigma^2_{v,t}} \right)^2 \cdot \left\{ \text{Cov} \left[ a_{t+1} + \theta_t s_{t+1}, \widetilde{V}_{t,1} (\mathbf{x}_{t+1}, a_{t+1}) \big| \mathcal{F}_t \right] + \text{Cov} \left[ \theta_t a_{t+1} + \theta_t s_{t+1}, \widetilde{V}_{t,2} (\mathbf{x}_{t+1}, s_{t+1}) \big| \mathcal{F}_t \right] - \text{Var} [a_{t+1} + \theta_t s_{t+1}] \text{E} [V(\mathbf{x}_{t+1})] \right\}
\]

which allows the application of the multivariate version of Stein’s Lemma. First, consider

\[
\text{Cov} \left[ a_{t+1} + \theta_t s_{t+1}, \widetilde{V}_{t,1} (\mathbf{x}_{t+1}, a_{t+1}) \big| \mathcal{F}_t \right] = \text{Cov} \left[ a_{t+1} + \theta_t s_{t+1}, (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})' \big| \mathcal{F}_t \right] \cdot E \left[ \frac{\partial \left( \mathbf{x}_{t+1} \atop a_{t+1} \atop s_{t+1} \atop \varepsilon_{z,t+1} \right)'}{\partial (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})'} \bigg| \mathcal{F}_t \right]
\]

where

\[
\frac{\partial \left( \mathbf{x}_{t+1} \atop a_{t+1} \atop s_{t+1} \atop \varepsilon_{z,t+1} \right)'}{\partial (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})'} =
\begin{pmatrix}
1 & \frac{\phi_t \hat{\sigma}^2_t}{\hat{\sigma}^2_t + \sigma^2_{v,t} + \sigma^2_{z,t+1}} & 0 & 0 & 1 \\
0 & \frac{\phi_t \hat{\sigma}^2_t}{\hat{\sigma}^2_t + \sigma^2_{v,t} + \sigma^2_{z,t+1}} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \sigma_z
\end{pmatrix}
\]

and

\[
\text{Cov} \left[ a_{t+1} + \theta_t s_{t+1}, (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})' \big| \mathcal{F}_t \right] = (\hat{\sigma}^2_t + \sigma^2_{v,t} + \theta_t \hat{\sigma}_t^2) \begin{pmatrix}
1 \\
1 \\
0
\end{pmatrix}
\]

and

\[
\text{Var} [a_{t+1} + \theta_t s_{t+1}] = \hat{\sigma}^2_t (1 + \theta_t) + \sigma^2_{v,t}
\]
Hence,\
\[ \text{Cov} \left[ a_{t+1} + t_t s_{t+1}, \hat{V}_{t,1} (x_{t+1}, a_{t+1}) \right] = (\hat{\sigma}^2_t + \sigma^2_{v,t} + t_t \hat{\sigma}^2_t) \cdot \right. \]
\[ \left. \left\{ E \left[ \frac{\partial V (x_{t+1})}{\partial y_{t+1}} (a_{t+1} - E [a_{t+1} | \mathcal{F}_t]) \right] \right| \mathcal{F}_t \right) + E \left[ V (x_{t+1}) \right| \mathcal{F}_t] + \phi_\mu \frac{\hat{\sigma}^2_t + \sigma^2_{v,t}}{\hat{\sigma}^2_{v,t}} E \left[ \frac{\partial V (x_{t+1})}{\hat{\mu}_{t+1}} (a_{t+1} - E [a_{t+1} | \mathcal{F}_t]) \right| \mathcal{F}_t \right) \}\]

Further, given that \( \frac{\partial V (x_{t+1})}{\partial y_{t+1}} \) satisfies condition (1) one may apply Lemma (6) to obtain
\[ \text{Cov} \left[ a_{t+1}, \frac{\partial V (x_{t+1})}{\partial \hat{\mu}_{t+1}} \right] = (\hat{\sigma}^2_t + \sigma^2_{v,t} + t_t \hat{\sigma}^2_t) \cdot \right. \]
\[ \left. \left\{ E \left[ \frac{\partial V (x_{t+1})}{\partial y_{t+1}} a_{t+1} \right] \right| \mathcal{F}_t \right) + E \left[ V (x_{t+1}) \right| \mathcal{F}_t] + \phi_\mu \frac{\hat{\sigma}^2_t + \sigma^2_{v,t}}{\hat{\sigma}^2_{v,t}} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \hat{\mu}_{t+1} \partial y_{t+1}} \right| \mathcal{F}_t \right) \right\} \]

Substituting this result yields:
\[ \text{Cov} \left[ a_{t+1} + t_t s_{t+1}, \hat{V}_{t,1} (x_{t+1}, a_{t+1}) \right] = (\hat{\sigma}^2_t + \sigma^2_{v,t} + t_t \hat{\sigma}^2_t) \cdot \right. \]
\[ \left. \left\{ E \left[ \frac{\partial V (x_{t+1})}{\partial y_{t+1}} a_{t+1} \right] \right| \mathcal{F}_t \right) + E \left[ V (x_{t+1}) \right| \mathcal{F}_t] + \phi_\mu \frac{\hat{\sigma}^2_t + \sigma^2_{v,t}}{\hat{\sigma}^2_{v,t}} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \hat{\mu}_{t+1} \partial y_{t+1}} \right| \mathcal{F}_t \right) \right\} \]

Next, one may write
\[ \text{Cov} \left[ t_t a_{t+1} + t_t s_{t+1}, \hat{V}_{t,2} (x_{t+1}, s_{t+1}) \right] = \text{Cov} \left[ t_t a_{t+1} + t_t s_{t+1}, (a_{t+1}, s_{t+1}, \xi_{t,t+1})' \right| \mathcal{F}_t \right] \cdot \right. \]
\[ \left. E \left[ d \left( \begin{array}{c} x_{t+1} \\ s_{t+1} \end{array} \right)' \right| \mathcal{F}_t \right) E \left[ \frac{\partial \hat{V}_{t,2} (x_{t+1}, s_{t+1})}{\partial \left( \begin{array}{c} x_{t+1} \\ s_{t+1} \end{array} \right)} \right| \mathcal{F}_t \right] \]

Note that
\[ \frac{\partial \left( \begin{array}{c} x_{t+1} \\ s_{t+1} \end{array} \right)'}{\partial \left( a_{t+1}, s_{t+1}, \xi_{t,t+1} \right)'} = \left( \begin{array}{cccc} 1 & \frac{\phi_\mu \hat{\sigma}^2_t}{\hat{\sigma}^2_t + \sigma^2_{v,t} + t_t \hat{\sigma}^2_t} & 0 & 0 & 0 \\ \frac{\phi_\mu \hat{\sigma}^2_t}{\hat{\sigma}^2_t + \sigma^2_{v,t} + t_t \hat{\sigma}^2_t} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \sigma_z \end{array} \right) \right) \]
and

\[
E \left[ \frac{\partial \tilde{V}_{t,2} (x_{t+1}, s_{t+1})}{\partial (x_{t+1})} \right] \mathcal{F}_t
\]

\[
= E \left[ \left( \frac{\partial V (x_{t+1})}{\partial x_{t+1}} \frac{\partial (x_{t+1})}{\partial s_{t+1}} \right) \right] \mathcal{F}_t
\]

and

\[
Cov \left[ \iota_t a_{t+1} + \iota_t s_{t+1}, (a_{t+1}, s_{t+1}, z_{t+1})' \right] \mathcal{F}_t
\]

\[
= \iota_t \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 + \iota_t \hat{\sigma}_t^2 \right) \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}^\prime
\]

Therefore, one obtains

\[
Cov \left[ \iota_t a_{t+1} + \iota_t s_{t+1}, \tilde{V}_{t,2} (x_{t+1}, s_{t+1}) \right] \mathcal{F}_t
\]

\[
= \iota_t \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 + \iota_t \hat{\sigma}_t^2 \right) \cdot \\
\left\{ Cov \left[ \frac{\partial V (x_{t+1})}{\partial y_{t+1}}, s_{t+1} \right] \mathcal{F}_t \right\} + \\
\phi \mu \hat{\sigma}_t^2 \left( 1 + \iota_t \right) \hat{\sigma}_t^2 \cdot E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}^2} \right] \mathcal{F}_t
\]

Further, given that \( \frac{\partial V (x_{t+1})}{\partial \mu_{t+1}} \) satisfies condition (1) one may apply Lemma (7) to obtain

\[
Cov \left[ s_{t+1}, \frac{\partial V (x_{t+1})}{\partial \mu_{t+1}} \right] \mathcal{F}_t
\]

\[
= \hat{\sigma}_t^2 E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}^2} \right] \mathcal{F}_t + \phi \mu \hat{\sigma}_t^2 \left( 1 + \iota_t \right) \hat{\sigma}_t^2 \cdot E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1} \partial y_{t+1}} \right] \mathcal{F}_t
\]

This yields

\[
Cov \left[ \iota_t a_{t+1} + \iota_t s_{t+1}, \tilde{V}_{t,2} (x_{t+1}, s_{t+1}) \right] \mathcal{F}_t
\]

\[
= \iota_t \left( \hat{\sigma}_t^2 + \sigma_{v,t}^2 + \iota_t \hat{\sigma}_t^2 \right) \cdot \\
\left\{ Cov \left[ \frac{\partial V (x_{t+1})}{\partial y_{t+1}}, s_{t+1} \right] \mathcal{F}_t + E \left[ V (x_{t+1}) \right] \mathcal{F}_t \right\} + \\
\phi \mu \hat{\sigma}_t^2 \left( 1 + \iota_t \right) \hat{\sigma}_t^2 \cdot E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}^2} \right] \mathcal{F}_t + \\
\phi \mu \hat{\sigma}_t^2 \left( 1 + \iota_t \right) \hat{\sigma}_t^2 \cdot E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1} \partial y_{t+1}} \right] \mathcal{F}_t
\]
Substituting the results into,

\[
E \left[ V(x_{t+1}) \frac{\partial f(a_{t+1}, s_{t+1} | x_t, u_{t})}{\partial \sigma_{p,t}} \right]_{F_t}
\]

\[
= \frac{1}{2} \frac{1}{(\hat{\sigma}^2_t (1 + \mu_t) + \sigma^2_{r,t})^2} \cdot 
\left\{ Cov \left[ a_{t+1} + \mu_t s_{t+1}, V_{t,1} (x_{t+1}, a_{t+1}) \right]_{F_t} + 
Cov \left[ \mu_t a_{t+1} + \mu_t s_{t+1}, V_{t,2} (x_{t+1}, s_{t+1}) \right]_{F_t} - 
Var \left[ a_{t+1} + \mu_t s_{t+1} \right]_{F_t} E \left[ V (x_{t+1}) \right]_{F_t} \right\}
\]

yields

\[
E \left[ V(x_{t+1}) \frac{\partial f(a_{t+1}, s_{t+1} | x_t, u_{t})}{\partial \sigma_{p,t}} \right]_{F_t}
\]

\[
= \frac{1}{2} \frac{1}{(\hat{\sigma}^2_t (1 + \mu_t) + \sigma^2_{r,t})^2} \cdot 
\left\{ \left( \hat{\sigma}^2_t + \sigma^2_{r,t} + \mu_t \hat{\sigma}^2_t \right) Cov \left[ \frac{\partial V (x_{t+1})}{\partial y_{t+1}}, a_{t+1} + \mu_t s_{t+1} \right]_{F_t} + 
(1 + \mu_t) \left( \hat{\sigma}^2_t + \sigma^2_{r,t} + \mu_t \hat{\sigma}^2_t \right) E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1} \partial y_{t+1}} \right]_{F_t} + 
\phi \mu \hat{\sigma}^2_t (1 + \mu_t) \left( \hat{\sigma}^2_t + \sigma^2_{r,t} + \hat{\sigma}^2_t \right) E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}^2} \right]_{F_t} - 
\phi \mu \hat{\sigma}^2_t (1 + \mu_t)^2 E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}^2} \right]_{F_t} \right\}
\]

Finally, given that \( \frac{\partial V(x_{t+1})}{\partial y_{t+1}} \) satisfies condition \( \text{(1)} \) one may apply Lemmas \( \text{(6)} \) and \( \text{(7)} \) to obtain

\[
Cov \left[ \left( a_{t+1} + \mu_t s_{t+1}, \frac{\partial V (x_{t+1})}{\partial y_{t+1}} \right) \right]_{F_t}
\]

\[
= \left( \hat{\sigma}^2_t + \sigma^2_{r,t} + \mu_t \hat{\sigma}^2_t \right) E \left[ \frac{\partial^2 V (x_{t+1})}{\partial y_{t+1}^2} \right]_{F_t} + 
\phi \mu \hat{\sigma}^2_t (1 + \mu_t) E \left[ \frac{\partial^2 V (x_{t+1})}{\partial y_{t+1} \partial \mu_{t+1}} \right]_{F_t}
\]

Using this result and

\[
Var \left[ a_{t+1} + \mu_t s_{t+1} \right]_{F_t}
\]

\[
= (1 + \mu_t) \left( \mu_t \hat{\sigma}^2_t + \sigma^2_{r,t} + \hat{\sigma}^2_t \right)
\]

yields the stated equation. \( \square \)
Lemma 14.

\[
E \left[ \frac{\partial \mu_{t+1}}{\partial \hat{\sigma}^2_{p,t}} \frac{\partial V (x_{t+1})}{\partial \mu_{t+1}} \right] = \phi \mu \sigma^2_{v,t} \left( \hat{\sigma}^2_t + \hat{\sigma}^2_t \iota_t + \sigma^2_{v,t} \right)
\]

Proof. Recall the previously derived formula

\[
\frac{\partial \hat{\mu}_{t+1}}{\partial \hat{\sigma}^2_{p,t}} = \phi \mu (a_t + (\hat{\mu}_t + \omega_{a,t})) + \iota_t (s_t + \hat{\mu}_t - \mu_t) \left( \hat{\sigma}^2_t + \hat{\sigma}^2_t \iota_t + \sigma^2_{v,t} \right)
\]

Thus, one may write

\[
E \left[ \frac{\partial \mu_{t+1}}{\partial \hat{\sigma}^2_{p,t}} \frac{\partial V (x_{t+1})}{\partial \mu_{t+1}} \right] = \phi \mu \sigma^2_{v,t} \left( \hat{\sigma}^2_t + \hat{\sigma}^2_t \iota_t + \sigma^2_{v,t} \right)
\]

Given that \( \frac{\partial V (x_{t+1})}{\partial \mu_{t+1}} \) satisfies condition (1) one may apply Lemmas (6) and (7) to obtain the stated relation.

Lemma 15.

\[
\frac{\partial V (x_t)}{\partial \hat{\sigma}^2_{p,t}} = \beta \phi \mu \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}} \frac{\partial V (x_{t+1})}{\partial y_{t+1}} \right] + \beta \phi \mu \left( 1 + \iota_t \right) \frac{\sigma^2_{v,t}}{(\hat{\sigma}^2_t + \hat{\sigma}^2_t \iota_t + \sigma^2_{v,t})^2} \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}} \frac{\partial V (x_{t+1})}{\partial y_{t+1}} \right]
\]

\[
\beta \phi \mu \left( 1 + \iota_t \right) \frac{\sigma^2_{v,t}}{(\hat{\sigma}^2_t + \hat{\sigma}^2_t \iota_t + \sigma^2_{v,t})^2} \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}} \frac{\partial V (x_{t+1})}{\partial y_{t+1}} \right] + \beta \phi \mu \left( 1 + \iota_t \right) \frac{\sigma^2_{v,t}}{(\hat{\sigma}^2_t + \hat{\sigma}^2_t \iota_t + \sigma^2_{v,t})^2} \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}} \frac{\partial V (x_{t+1})}{\partial y_{t+1}} \right] + \beta \phi \mu \left( 1 + \iota_t \right) \frac{\sigma^2_{v,t}}{(\hat{\sigma}^2_t + \hat{\sigma}^2_t \iota_t + \sigma^2_{v,t})^2} \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}} \frac{\partial V (x_{t+1})}{\partial y_{t+1}} \right]
\]
Proof. By the envelope theorem one may write

\[
\frac{\partial V (x_t)}{\partial \hat{\sigma}^2_{p,t}} = \beta_t E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t}} \right] + \beta_t E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t}} \right] f(\sigma_{t+1}, \hat{\sigma}^2_{p,t}, \sigma^2_{v,t}, \hat{\sigma}^2_{p,t}) \bigg| \mathcal{F}_t \right] \\
= \beta_t E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t}} \right] \frac{f(\sigma_{t+1}, \hat{\sigma}^2_{p,t}, \sigma^2_{v,t}, \hat{\sigma}^2_{p,t})}{f(\sigma_{t+1}, \hat{\sigma}^2_{p,t}, \sigma^2_{v,t}, \hat{\sigma}^2_{p,t})} \bigg| \mathcal{F}_t \right] \\
= \beta_t E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t}} \right] f(\sigma_{t+1}, \hat{\sigma}^2_{p,t}, \sigma^2_{v,t}, \hat{\sigma}^2_{p,t}) \bigg| \mathcal{F}_t \right] \\
= \beta_t E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t}} \right] f(\sigma_{t+1}, \hat{\sigma}^2_{p,t}, \sigma^2_{v,t}, \hat{\sigma}^2_{p,t}) \bigg| \mathcal{F}_t \right] \\
= 0
\]

since

\[
\beta_t E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t}} \right] f(\sigma_{t+1}, \hat{\sigma}^2_{p,t}, \sigma^2_{v,t}, \hat{\sigma}^2_{p,t}) \bigg| \mathcal{F}_t \right] \\
= \beta_t E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t}} \right] f(\sigma_{t+1}, \hat{\sigma}^2_{p,t}, \sigma^2_{v,t}, \hat{\sigma}^2_{p,t}) \bigg| \mathcal{F}_t \right] \\
= 0
\]

where

\[
\frac{\partial f(a_{t+1}, s_{t+1} | x_t, u_t)}{\partial \sigma^2_{p,t}} \frac{f(a_{t+1}, s_{t+1} | x_t, u_t)}{f(a_{t+1}, s_{t+1} | x_t, u_t)} = \frac{1}{2} \frac{\text{Var} [a_{t+1} + \epsilon t s_{t+1} | \mathcal{F}_t]}{(\sigma^2_t + \sigma^2_{v,t})^2}
\]

as defined and derived above. Substituting the equations provided by Lemma (14) and (13) and gathering common terms yields the stated equation. \[ \square \]

Lemma 16.

\[
E_t \left[ V (x_{t+1}) \frac{\partial f(a_{t+1}, s_{t+1} | x_t, u_t)}{\partial \sigma^2_{p,t}} f(a_{t+1}, s_{t+1} | \mathcal{F}_t) \right] = \frac{1}{2} \frac{\partial \sigma^2_{p,t+1}}{\partial \sigma_t} E_t \left[ \frac{\partial^2 V (x_{t+1})}{\partial \sigma^2_{p,t+1}} \right] (48)
\]
Proof. Based on the previous definitions one may write

\[
\frac{\partial f(a_{t+1}, s_{t+1} | x_t, u_t)}{\partial s_t} \tilde{\Phi}_{x_t, u_t} = \frac{\partial^2 f(a_{t+1}, s_{t+1} | x_t, u_t)}{2} \frac{\partial s_t^2}{\partial \sigma_{v,v,t}^2} \tilde{\Phi}_{s_t, u_t+1}
\]

\[
\frac{\partial f(a_{t+1}, s_{t+1} | x_t, u_t)}{\partial u_t} = \frac{\partial^2 f(a_{t+1}, s_{t+1} | x_t, u_t)}{2} \frac{\partial u_t^2}{\partial \sigma_{v,v,t}^2} \tilde{\Phi}_{u_t, s_t+1}
\]

Applying Lemma (8) directly yields the stated equation. \( \square \)

A.4 Propositions

Proposition 5. The decision problem associated with the Bellman equation (27) implies the following necessary intertemporal condition for optimal information acquisition \( \sigma_t \):

\[
E_t \left[ m_{t+1} + \delta_t + MC_t (\sigma^2_{p,t+1}) \right] = 1
\]

where \( m_{t+1} = \frac{\beta_t F'(c_{t+1})}{F'(c_t)} \) denotes the stochastic discount factor and the following definitions apply

\[
\Pi_t = \frac{\partial V^{RS}(x_{t+1}, u(x_{t+1})) \partial k_{t+1}(x_{t+1})}{\partial \sigma_{v,v,t}^2} \tilde{\Phi}_{V^{RS}(x_{t+1}, u(x_{t+1)))}
\]

\[
MC_t (\sigma^2_{p,t+1}) = n_t' (t_t) \left( \frac{\partial^2 V^{RS}(x_{t+1}, u(x_{t+1))) \partial k_{t+1}(x_{t+1})}{\partial \sigma_{v,v,t}^2} \tilde{\Phi}_{V^{RS}(x_{t+1}, u(x_{t+1)))}
\]

\[
\frac{\partial^2 \sigma_{p,t+1} (\sigma^2_{p,t+1})}{\partial \sigma_{v,v,t}^2} = \frac{\partial^2 V^{RS}(x_{t+1}, u(x_{t+1))) \partial k_{t+1}(x_{t+1})}{\partial \sigma_{v,v,t}^2} \tilde{\Phi}_{V^{RS}(x_{t+1}, u(x_{t+1)))}
\]

and where \( V^{RS}(x_{t+1}, u(x_{t+1})) \) denotes the right side of the Bellman equation (27) evaluated at the optimal policy rule \( u(x_{t+1}) \).

Proof. Note that the first-order necessary condition for \( \sigma_t \), as derived above,
may be written as follows:

\[
E \left[ \frac{\partial \hat{\mu}_{t+1}}{\partial \mu_t} \frac{\partial V (x_{t+1})}{\partial \hat{\mu}_{t+1}} \right] F_t + E \left[ \frac{\partial \hat{\sigma}^2_{p,t+1}^2}{\partial \hat{\sigma}^2_{p,t+1}} \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t + (49)
\]

\[
E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t
\]

\[
= \frac{F'(c_t) n_t'(u_t)}{\beta_t}
\]

Substituting the previous results (32) and (48), one obtains

\[
- \frac{\partial \hat{\sigma}^2_{p,t+1}}{\partial \hat{\sigma}^2_{p,t+1}} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t + \frac{\partial \hat{\sigma}^2_{p,t+1}}{\partial \hat{\sigma}^2_{p,t+1}} E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t + \frac{1}{2} \frac{\partial \hat{\sigma}^2_{p,t+1}}{\partial \hat{\sigma}^2_{p,t+1}} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t
\]

\[
= \frac{F'(c_t) n_t'(u_t)}{\beta_t}
\]

or

\[
E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t = \frac{1}{2} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t + \frac{F'(c_t) n_t'(u_t)}{\partial \hat{\sigma}^2_{p,t+1}} \beta_t (50)
\]

Define

\[
\theta_{t,\hat{\sigma}^2_{p,t}} = \frac{1}{2} \beta_t \cdot E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t + \beta_t \phi \mu E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t
\]

\[
\delta_t = \frac{\partial \hat{\sigma}^2_{p,t+1}}{\partial \hat{\sigma}^2_{p,t}} = \frac{\partial \hat{\sigma}^2_{t+1}}{\partial \hat{\sigma}^2_{t}}
\]

\[
\gamma_{t,\hat{\sigma}^2_{p,t}} = \beta_t \phi^2 \mu (1 + \hat{\mu}_t) \hat{\sigma}^2_{v,t} + \frac{1}{2} (1 + \hat{\mu}_t) \hat{\sigma}^2_{v,t} (\hat{\sigma}^2_{t} + \hat{\sigma}^2_{v,t})
\]

such that by (47) \( \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \) may be written as follows

\[
\frac{\partial V (x_t)}{\partial \hat{\sigma}^2_{p,t}} = \theta_{t,\hat{\sigma}^2_{p,t}} + \beta_t \delta_t E \left[ \frac{\partial V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t + \gamma_{t,\hat{\sigma}^2_{p,t}} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \hat{\sigma}^2_{p,t+1}} \right] F_t \).
Substituting (50), rearranging, and going one period forward gives

\[
\frac{\partial V (x_{t+1})}{\partial \sigma^2_{p,t+1}} = \theta_{t+1} \delta_{p,t+1} + \beta_{t+1} \delta_{t+1} + \frac{F' (c_{t+1}) n'_{t+1} (u_{t+1})}{\beta_{t+1} \hat{\sigma}^2_{p,t+1} + 2} + \left( \gamma_{t+1} \delta_{p,t+1} + \frac{1}{2} \beta_{t+1} \delta_{t+1} \right) E \left[ \frac{\partial^2 V (x_{t+2})}{\partial \mu^2_{t+2}} \bigg| F_{t+1} \right]
\]

Using the fact that

\[
\gamma_{t+1} \delta_{p,t+1} + \frac{1}{2} \beta_{t+1} \delta_{t+1} = \frac{1}{2} \sigma^2_{p,t+1}
\]

and taking mathematical expectations with respect to \( F_t \) yields

\[
E \left[ \frac{\partial V (x_{t+1})}{\partial \sigma^2_{p,t+1}} \bigg| F_t \right] = E \left[ \theta_{t+1} \delta_{p,t+1} \bigg| F_t \right] + E \left[ \beta_{t+1} \delta_{t+1} + \frac{F' (c_{t+1}) n'_{t+1} (u_{t+1})}{\beta_{t+1} \hat{\sigma}^2_{p,t+1} + 2} \bigg| F_t \right] + \frac{1}{2} \sigma^2_{p,t+1} \beta_{t+1} E \left[ \frac{\partial^2 V (x_{t+2})}{\partial \mu^2_{t+2}} \bigg| F_t \right] - \frac{1}{2} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu^2_{t+1}} \bigg| F_t \right]
\]

Substituting (51) into (50) and rearranging yields

\[
E \left[ \frac{\partial^2 V (x_{t+2})}{\partial \mu^2_{t+2}} \bigg| F_t \right] - \frac{1}{2} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu^2_{t+1}} \bigg| F_t \right]
\]

From equation (33) it follows that

\[
\frac{1}{2} E \left[ \phi_{\mu} \beta_{t+1} E \left[ \frac{\partial^2 V (x_{t+2})}{\partial \mu^2_{t+2}} \bigg| F_t \right] - \frac{\partial^2 V (x_{t+1})}{\partial \mu^2_{t+1}} \bigg| F_t \right] - \beta_{t+1} \phi_{\mu} E \left[ \frac{\partial^2 V (x_{t+2})}{\partial \mu^2_{t+2}} \bigg| F_{t+1} \right] - \frac{1}{2} E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu^2_{t+1}} - \frac{\partial^2 V_{RS}(x_{t+1}, u (x_{t+1}))}{\partial \mu^2_{t+1}} \bigg| F_t \right]
\]

51
Substituting (53) and the definition of $\theta_{t+1},\hat{\sigma}_{p,t+1}^2$ into (52) and rearranging yields the result

$$E\left[\delta_{t+1}\left(\frac{\partial^2 \hat{\sigma}_{p,t+1}^2}{\partial \mu_1^2} n'_t (t+1) \beta_1 F'(c_{t+1})}{\partial \mu_1^2} F'(c_t) \right| \mathcal{F}_t\right] +$$

$$\frac{1}{2} \beta_t E \left[ \frac{\partial^2 V(x_{t+1})}{\partial \mu_1^2} - \frac{\partial^2 V^{RS}(x_{t+1}, u(x_{t+1}))}{\partial \mu_1^2} \right| \mathcal{F}_t\right]$$

$$= 1$$

where

$$= \frac{\partial^2 V (x_{t+1})}{\partial \mu_1^2} - \frac{\partial^2 V^{RS}(x_{t+1}, u(x_{t+1}))}{\partial \mu_1^2}$$

$$= 2 \frac{\partial^2 V^{RS}(x_{t+1}, u(x_{t+1}))}{\partial \mu_1^2} \frac{\partial u (x_{t+1})}{\partial \mu_1}$$

First order optimality ensures that

$$\frac{\partial V^{RS}(x_{t+1}, u(x_{t+1}))}{\partial u} = \left( \begin{array}{c} 0 \\ 0 \end{array} \right).$$

Differentiating w.r.t. $\mu_{t+1}$ yields

$$\frac{\partial V^{RS}(x_{t+1}, u(x_{t+1}))}{\partial u(x_{t+1})} \frac{\partial u(x_{t+1})}{\partial \mu_{t+1}} = 0$$

or

$$\frac{\partial u(x_{t+1})}{\partial \mu_{t+1}} = - \left( \frac{\partial V^{RS}(x_{t+1}, u(x_{t+1}))}{\partial u} \right)^{-1} \frac{\partial^2 V^{RS}(x_{t+1}, u(x_{t+1}))}{\partial u(x_{t+1})} \frac{\partial \mu_{t+1}}{\partial \mu_{t+1}}.$$

52
This yields
\[
\begin{align*}
\frac{\partial^2 V (x_{t+1})}{\partial \hat{\mu}_{t+1}^2} &= \frac{\partial^2 V^{RS} (x_{t+1}, u (x_{t+1}))}{\partial \hat{\mu}_{t+1}^2} \\
&= 2 \frac{\partial^2 V^{RS} (x_{t+1}, u (x_{t+1}))}{\partial \hat{\mu}_{t+1}^2} \frac{\partial u (x_{t+1})}{\partial \hat{\mu}_{t+1}} \\
&\quad + \frac{\partial^2 V^{RS} (x_{t+1}, u (x_{t+1}))}{\partial \hat{\mu}_{t+1}^2} \left( \frac{\partial u (x_{t+1})}{\partial \hat{\mu}_{t+1}} \right)^2 \\
&= \left( \frac{\partial^2 V^{RS} (x_{t+1}, u (x_{t+1}))}{\partial \hat{\mu}_{t+1}^2} \right) \frac{\partial u (x_{t+1})}{\partial \hat{\mu}_{t+1}} + \frac{\partial^2 V^{RS} (x_{t+1}, u (x_{t+1}))}{\partial \hat{\mu}_{t+1}^2} \left( \frac{\partial u (x_{t+1})}{\partial \hat{\mu}_{t+1}} \right). \\
\end{align*}
\]

This yields the final result
\[
E_t \left[ m_{t, t+1} + \delta_{t+1} MC_{t+1} \left( \hat{\sigma}^2 p_{t, t+1} \right) \right] = 1
\]
where \( m_{t, t+1} = \frac{\hat{\delta}^2 V^{RS} (x_{t+1}, u_{t+1})}{F'(c_{t+1})} \) denotes the stochastic discount factor and the following definitions apply

\[
\Pi_{t+1} = \frac{1}{2} \frac{\partial^2 V^{RS} (x_{t+1}, u_{t+1})}{\partial \hat{\mu}_{t+1}^2} \frac{\partial k_{t+1} (x_{t+1})}{\partial \hat{\mu}_{t+1}} + \frac{\partial^2 V^{RS} (x_{t+1}, u_{t+1})}{\partial \hat{\mu}_{t+1}^2} \frac{\partial k_{t+1} (x_{t+1})}{\partial \hat{\mu}_{t+1}}
\]

\[
MC_t (\hat{\sigma}^2 p_{t, t+1}) = n_t^l (t_i) \sqrt{-\frac{\partial \hat{\sigma}^2 p_{t, t+1}}{\partial t}}
\]

\[
\frac{\partial \hat{\sigma}^2 p_{t, t+1}}{\partial t} = - \frac{\hat{\delta}^2 p_{t, t+1}}{(1 + t_i + \frac{\sigma^2_{t_i}}{\sigma^2_{t}})^2}
\]

\[
\delta_{t+1} = \frac{\partial \hat{\sigma}^2 p_{t, t+1}}{\partial t} = \frac{\hat{\delta}^2 p_{t}}{(1 + (1 + t_{i+1}) \frac{\sigma^2_{t+1}}{\sigma^2_{t}})^2}
\]

Note that given a well defined problem with a negative definite Hessian \( (\partial^2 V^{RS} / \partial u_{t+1}^2) \), the term \( \Pi_{t+1} \) is positive state-by-state.

\[
\square
\]

53
A first order approximation equal to zero, i.e.
as follows:

**Condition 3.**

**Definition 2.**

\[ b(x_t, \phi_{\mu}) = - \frac{\sigma_{v,t}^2 \beta_t}{\left(1 + \ell_t(x_t, \phi_{\mu}) + \sigma_{v,t}^2 \right)^2} \left\{ E_t \left[ \frac{\partial V(x_{t+1}(\phi_{\mu}), \phi_{\mu})}{\partial y_{t+1}} - \frac{1}{2} \frac{\partial^2 V(x_{t+1}(\phi_{\mu}), \phi_{\mu})}{\partial \mu t_{t+1}} \right] \right\} + E_t \left[ \phi_{\mu} \cdot \partial \frac{\partial V(x_{t+1}(\phi_{\mu}), \phi_{\mu})}{\partial \mu t_{t+1}} \partial \mu t_{t+1} \left( \frac{1}{2} \frac{\partial^2 V(x_{t+1}(\phi_{\mu}), \phi_{\mu})}{\partial \mu t_{t+1}} \right) \right] \]

\[ \mu_{t+1} = \frac{\partial^2 V_{RS}(x_{t+1}, u_{t+1})}{\partial \mu_{t+1} \partial \beta_{t+1}} \frac{\partial \mu_{t+1}(x_{t+1})}{\partial \beta_{t+1}} F'(c_{t+1}) \]

\[ \mu_{t+1} = - \frac{\phi_{\mu} \beta_{t+1} \sigma_{v,t+1}^2}{1 + \ell_{t+1} + \frac{\sigma_{v,t+1}^2}{\sigma_{t+1}^2}} \frac{\partial t_{t+1}(x_{t+1})}{\partial \mu_{t+1}} \]

\[ \left\{ E_{t+1} \left[ \partial \frac{\partial V(x_{t+2})}{\partial \beta_{t+2}} \partial y_{t+2} - \frac{1}{2} \frac{\partial^2 V(x_{t+2})}{\partial \beta_{t+2}} \right] \right\} + E_{t+1} \left[ \phi_{\mu} \cdot \partial \frac{\partial V(x_{t+2})}{\partial \beta_{t+2}} \partial \mu_{t+2} \left( \frac{1}{2} \frac{\partial^2 V(x_{t+2})}{\partial \beta_{t+2}} \right) \right] \]

**Proposition 6.** The addends that enter the period gain \( \Pi_{t+1} \) may be written as follows:

For small \( \phi_{\mu} \) and given condition (3) is satisfied \( \frac{\partial^2 V_{RS}(x_{t+1}, u_{t+1})}{\partial \mu_{t+1} \partial \beta_{t+1}} \) is to a first order approximation equal to zero, i.e.

\[ \approx \phi_{\mu}^2 b(x_{t+1}, \phi_{\mu}) \bigg|_{\phi_{\mu}=0} + \frac{d}{d\phi_{\mu}} \left\{ \phi_{\mu}^2 b(x_{t+1}, \phi_{\mu}) \right\} \bigg|_{\phi_{\mu}=0} \]

\[ = 0. \]
Proposition 7. In the case of a linear period return function the following first order approximation of $\Pi_{t+1}$ around $\phi_{\mu} = 0$ holds:

$$
\Pi_{t+1}(x_{t+1}, \phi_{\mu}) \approx \Pi_{t+1}(x_{t+1}, \phi_{\mu})|_{\phi_{\mu}=0} + \frac{\partial \Pi_{t+1}(x_{t+1}, \phi_{\mu})}{\partial \phi_{\mu}}|_{\phi_{\mu}=0} \phi_{\mu}
$$

$$
= \frac{k_{t+1}(x_{t+1})}{2\eta_{t+1}(x_{t+1})},
$$

where

$$
\eta_{t+1}(x_{t+1}) \equiv -\frac{G''(k_{t+1}(x_{t+1}))}{G'(k_{t+1}(x_{t+1}))} k_{t+1}(x_{t+1}).
$$

If, in addition, $G(k_{t}) = k_{t}^{1-\eta}$, with $0 < \eta < 1$, we obtain

$$
\Pi_{t+1} \approx \left( (1 - \eta) E_{t+1} [A_{t+2}] \beta_{t+1} \right)^{\frac{1}{\beta_{t+1}}},
$$

Proof of Propositions 2 & 3. To determine the vector $\frac{\partial V^{RS}(x_{t}, u_{t})}{\partial u_{t} \partial \mu_{t}}$ differentiate the following vector w.r.t. $\tilde{\mu}_{t}$:

$$
\frac{\partial V^{RS}(x_{t}, u_{t})}{\partial u_{t}} = \beta_{t} \left( \frac{G'(k_{t})}{G(k_{t})} E \left[ \frac{\partial V'(x_{t+1})}{\partial y_{t+1}} \right] F_{t} \right)
$$

First, consider the element $\frac{\partial^{2}V^{RS}(x_{t}, u_{t})}{\partial k_{t} \partial \mu_{t}}$:

$$
\frac{\partial^{2}V^{RS}(x_{t}, u_{t})}{\partial k_{t} \partial \mu_{t}} = \beta_{t} \frac{G'(k_{t})}{G(k_{t})} \partial \left[ E \left[ \frac{\partial V'(x_{t+1})}{\partial y_{t+1}} \right] F_{t} \right]
$$

which directly follows from equation (38). Further, given that $\frac{\partial V'(x_{t+1})}{\partial y_{t+1}}$ satisfies condition (1), one may obtain by Lemma (1)

$$
E \left[ \frac{(s_{t+1} - E[s_{t+1}|F_{t}]) \partial V'(x_{t+1})}{\partial y_{t+1}} \right] F_{t}
$$

$$
= E \left[ \frac{\partial^{2}V'(x_{t+1})}{\partial y_{t+1}^{2}} \right] F_{t} + \phi_{\mu} E \left[ \frac{\partial^{2}V'(x_{t+1})}{\partial \mu_{t+1} \partial y_{t+1}} \right] F_{t}
$$
Next, consider the element \( \frac{\partial^2 V_{RS}(x_t, u_t)}{\partial t \partial \tilde{\mu}_t} \):

\[
\frac{\partial^2 V_{RS}(x_t, u_t)}{\partial t \partial \tilde{\mu}_t} = \partial \left\{ \beta_t \frac{\sigma_{p,t+1}^2}{\tilde{\mu}_t} E \left[ \frac{\partial V(x_{t+1})}{\partial \tilde{\mu}_{t+1}} - \frac{1}{2} \frac{\partial^2 V(x_{t+1})}{\partial \tilde{\mu}_{t+1}^2} \right] | F_t \right\} + \\
\beta_t \frac{\partial^2 \sigma_{p,t+1}^2}{\partial t} E \left[ \frac{\partial V(x_{t+1})}{\partial \sigma_{p,t+1}^2} \frac{1}{2} \frac{\partial^2 V(x_{t+1})}{\partial \tilde{\mu}_{t+1}^2} \right] | F_t \left( \frac{\partial \mu \sigma_{p,t+1} | F_t}{\partial \tilde{\mu}_t} \right)
\]

where the last step follows from Lemma \([5]\) given that \( \frac{\partial^2 V(x_{t+1})}{\partial \sigma_{p,t+1}^2} \) and \( \frac{\partial^2 V(x_{t+1})}{\partial \tilde{\mu}_{t+1}^2} \) satisfy condition \([1]\). This yields

\[
\frac{\partial^2 V_{RS}(x_t, u_t)}{\partial t \partial \tilde{\mu}_t} = \beta_t \frac{\partial^2 \sigma_{p,t+1}^2}{\partial t} .
\]

\[
\left\{ E \left[ \frac{\partial^2 V(x_{t+1})}{\partial \sigma_{p,t+1}^2} \frac{1}{2} \frac{\partial^2 \sigma_{p,t+1}^2}{\partial \tilde{\mu}_t | F_t} \right] \right\} + \\
\phi_t E \left[ \frac{\partial^2 V(x_{t+1})}{\partial \sigma_{p,t+1}^2} \frac{1}{2} \frac{\partial^3 \sigma_{p,t+1}^2}{\partial \tilde{\mu}_t | F_t} \left( \frac{\partial \mu \sigma_{p,t+1} | F_t}{\partial \tilde{\mu}_t} \right) \right] \}
\]

and thus

\[
\frac{\partial^2 V_{RS}(x_t, u_t)}{\partial k_t \partial \tilde{\mu}_t} = \beta_t \frac{G_t'(k_t)}{G_t(k_t)} E \left[ \frac{(s_{t+1} - E[s_{t+1}|F_t])}{\tilde{\sigma}_t^2} \frac{\partial V(x_{t+1})}{\partial y_{t+1}} \right] F_t \left( \frac{\partial \mu \beta | F_t}{\partial \tilde{\mu}_t} \right)
\]

\[
= \beta_t \frac{G_t'(k_t)}{G_t(k_t)} E \left[ \frac{(s_{t+1} - E[s_{t+1}|F_t])}{\tilde{\sigma}_t^2} A_{t+1} G_t(k_t) F'(c_{t+1}) \right] F_t \left( \frac{\partial \mu \beta | F_t}{\partial \tilde{\mu}_t} \right)
\]

\[
= \frac{\text{Cov}[s_{t+1}, A_{t+1} G_t(k_t) \beta F'(c_{t+1})]}{\tilde{\sigma}_t^2}
\]

56
where the last step follows from Lemma (7) given that \( \frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} \) and \( \frac{\partial^2 V(x_{t+1})}{\partial \mu_{t+1}^2} \) satisfy condition [1]. We therefore may summarize the results

\[
\begin{align*}
\frac{\partial^2 V_{RS}(x_t, u_t)}{\partial \mu_t \partial x_t} \frac{\partial x_t}{\partial \mu_t} & = \frac{\text{Cov} \left( s_{t+1}, A_{t+1} G' (k_t) m_{t+1} \right)}{\text{Var}_t [\mu_{t+1}]} \frac{\partial k_t (x_t)}{\partial \mu_t} \\
\frac{\partial^2 V_{RS}(x_t, u_t)}{\partial \mu_t \partial x_t} & = -\frac{\phi_t^2 \beta_t \sigma_{v,t}^2}{F'' (c_t) \left( 1 + \eta_t + \frac{\sigma_{v,t}^2}{\eta_t^2} \right)^2} \frac{\partial x_t}{\partial \mu_t} \\
& + \left\{ E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}^2} - \frac{1}{2} \frac{\partial^3 V (x_{t+1})}{\partial \mu_{t+1}^3} \right] \right\} + \phi_t E \left[ \frac{\partial^2 V (x_{t+1})}{\partial \mu_{t+1}^2} - \frac{1}{2} \frac{\partial^3 V (x_{t+1})}{\partial \mu_{t+1}^3} \right].
\end{align*}
\]

Further, given condition [3] is satisfied, we may write a first order approximation of \( \frac{\partial^2 V_{RS}(x_t, u_t) \partial x_t}{\partial \mu_t} \) around \( \phi_t = 0 \) as follows:

\[
\begin{align*}
\frac{\partial^2 V_{RS}(x_t, u_t, \phi_t)}{\partial \mu_t \partial x_t} & \approx \left\{ \phi_t^2 b (x_t, \phi_t) \right\} \bigg|_{\phi_t = 0} + \frac{d \left\{ \phi_t^2 b (x_t, \phi_t) \right\}}{d \phi_t} \bigg|_{\phi_t = 0} \\
& = \left\{ \phi_t^2 b (x_t, \phi_t) \right\} \bigg|_{\phi_t = 0} + \left\{ \phi_t^2 \frac{\partial b (x_t, \phi_t)}{\partial \phi_t} + 2 \phi_t b (x_t, \phi_t) \right\} \bigg|_{\phi_t = 0} \\
& = 0.
\end{align*}
\]

Hence, to a first order approximation of \( \phi_t \) around \( \phi_t = 0 \), the term \( \frac{\partial^2 V_{RS}(x_t, u_t) \partial x_t}{\partial \mu_t} \) is negligible. Further, for a linear period return function, one may write:

\[
\begin{align*}
\frac{\text{Cov} \left( s_{t+1}, A_{t+1} G' (k_t) m_{t+1} \right)}{\text{Var}_t [\mu_{t+1}]} \frac{\partial k_t (x_t)}{\partial \mu_t} & = \frac{1}{2} \text{Cov} \left( s_{t+1}, A_{t+1} G' (k_t) \right) \beta_t \frac{\partial k_t (x_t)}{\partial \mu_t} \\
& + \frac{1}{2} \text{Var}_t [\mu_{t+1}] E_t [A_{t+1}] G' (k_t) \beta_t \frac{\partial k_t (x_t)}{\partial \mu_t} \\
& = \frac{1}{2} \text{Var}_t [\mu_{t+1}] E_t [A_{t+1}] G' (k_t) \beta_t \frac{\partial k_t (x_t)}{\partial \mu_t} \\
& = 1 \frac{\partial k_t (x_t)}{\partial \mu_t} \\
& = \frac{k_t (x_t)}{2 \eta_t}.
\end{align*}
\]
where
\[ \eta_t \equiv -\frac{G''(k_t)}{G'(k_t)} k_t, \]
and where we used the fact that by first order optimality the following relation holds,
\[ \beta_t E_t [A_{t+1}] G'(k_t) = 1, \]
with
\[ E_t [A_{t+1}] = \exp \left\{ \omega_{a} + \hat{\mu}_t + \frac{1}{2} \text{Var}_t [a_{t+1}] \right\}, \]
and which, by the implicit function theorem, yields the equation
\[ \frac{\partial k_t (x_t)}{\partial \hat{\mu}_t} = \frac{k_t}{\eta_t}. \]
Further, for \( G_t (k_t) = k_t^{1-\eta} \) one obtains
\[ k_t = \left( (1 - \eta) E_t [A_{t+1}] \beta_t \right)^{\frac{1}{\eta}} \]
\[ \eta_t = \eta \]
and thus
\[ \Pi_t = \frac{\left( (1 - \eta) E_t [A_{t+1}] \beta_t \right)^{\frac{1}{\eta}}}{2 \eta}. \]
This completes the proof.

**Proposition 8 (Human Capital).** The decision problem associated with the Bellman equation (27), adjusted by the specifications
\[ \mu_t = \sigma^2_{\mu,t-1} \varepsilon_{\mu,t} \]
\[ a_{t+1} = \omega_{a,t} + \mu_t + \sigma_{\varepsilon_{a,t+1}} \]
\[ \omega_{a,t+1} = \bar{\omega} + \phi_{\omega} \omega_{a,t} + \sigma_{\omega} \varepsilon_{\omega,r+1} \]
\[ h_t = \delta h_{t-1} \text{ with } 0 < \delta_h < 1 \]
\[ n (t_t, h_t) = \chi \frac{(t_t - h_t)^2}{2} \text{ with } \chi > 0, \]
implies the following necessary intertemporal condition for optimal information acquisition \( t \) :
\[ E_t \left[ m_{t+1} \Pi_{t+1} + \delta_{t+1} MC_{t+1} \right] = 1 \]
where \( m_{t+1} = \frac{\beta_t F'(c_{t+1})}{F'(c_t)} \) denotes the stochastic discount factor, and

\[
\Pi_{t+1} = \frac{1}{2} \left( \frac{s_{t+1} G(k_{t+1}) - Cov_{t+1} [a_{t+2}, m_{t+2} Y_{t+2}]}{Var_{t+1} [a_{t+2}]} \right)
\]

\[
MC_t = -n_t (\mu_t, h_t) / \partial \delta_{t+1}^2 / \partial t_t
\]

\[
\sigma_{t+1} = \delta_{h} \partial \sigma^2_{t+2} / \partial t_{t+1} / \partial \delta_{t+1}^2 / \partial t_t
\]

\[
\partial \delta^2_{t+1} / \partial t_t = -\sigma_{\mu,t}^2 / \sigma^2_{v,t} \left( 1 + \mu_t \frac{\sigma^2_{\mu,t}}{\sigma^2_{v,t}} \right)^{-2}
\]

and where \( \sigma^2_{t+1} = \sigma^2_{\mu,t, t+1} \). Under log-utility \( E_t \left[ \Pi_{t+1}^{1 + \eta} \right] = 0 \), implying steady state information acquisition equal to zero \( (i = 0) \). Under linear utility with exogenous, \( F_t \)-measurable discount factors \( m_{t+1} = \beta_t \), and \( G(k_t) = k^{1-\eta} \)

with \( 0 < \eta < 1 \), the following relation holds

\[
E_t \left[ \Pi_{t+1}^{\eta} \right] = E_t \left[ \frac{(1 - \eta) m_{t+2} E_t [A_{t+2}]}{2\eta} \right]
\]

Further, if \( m_{t+2} \) is \( F_t \) measurable, we obtain

\[
E_t \left[ \Pi_{t+1}^{\eta} = 0 \right] = \frac{(1 - \eta) m_{t+2} \frac{\hat{a}}{2\eta} \left( \hat{\omega} + \phi \omega_{u,t} + \frac{1}{2} \hat{\sigma}_{t+1}^2 + \frac{1}{2} \hat{\sigma}_{t+2}^2 + \frac{1}{2} \hat{\sigma}_{t+3}^2 + \frac{1}{2} \hat{\sigma}_{t+4}^2 + \frac{1}{2} \hat{\sigma}_{t+5}^2 \right) + \hat{\omega} \hat{\sigma}_{t+1}^2 / Var_t [\mu_t] \right]
\]

**Proof.** In this setup, the state is characterized by the vector \( x_t = (y_t, h_t, \mu_t, \sigma_t^2, z_t)' \), where \( \mu_t = E_t [\mu_t] \) and \( \sigma_t^2 = Var_t [\mu_t] \). The law of motion is given by:

\[
\begin{align*}
y_{t+1} &= a_{t+1} + \log G(k_t) \\
h_{t+1} &= \delta h_t \\
\mu_{t+1} &= \frac{\sigma_{\mu,t}^2}{\sigma_{\mu,t}^2 + \sigma_{t+1}^2} s_{t+1} \\
\sigma_{t+1}^2 &= \frac{\sigma_{\mu,t}^2 + \sigma_{t+1}^2}{1 + \sigma_{\mu,t}^2} \\
z_{t+1} &= \omega z + \phi z t + \sigma z \varepsilon_{z,t+1}
\end{align*}
\]

The first order necessary condition for \( \omega_t \) is given by

\[
F' (c_t) n_t (\omega_t, h_t)
\]

\[
= \int \beta_t \left( \frac{\partial \mu_{t+1}}{\partial t_t} \frac{\partial V (x_{t+1})}{\partial t_t} + \frac{\partial \sigma_{t+1}^2}{\partial t_t} \frac{\partial V (x_{t+1})}{\partial \sigma_{t+1}^2} + \frac{\partial h_{t+1}}{\partial h_t} \frac{\partial V (x_{t+1})}{\partial h_{t+1}} \right)
\]

\[
\int f (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1} | x_t, u_t) d \lambda (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1}) + \\
\int \beta_t V (x_{t+1}) \frac{\partial f (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1} | x_t, u_t)}{\partial t_t} d \lambda (a_{t+1}, s_{t+1}, \varepsilon_{z,t+1})
\]

59
Further, using the envelope theorem

\[ F'(c_t) n_t (\mu_t, h_t) \]

\[ = E_t \left[ \beta_t \left( \frac{\partial \hat{h}_{t+1}}{\partial \mu_{t+1}} \frac{\partial V (x_{t+1})}{\partial h_{t+1}} + \frac{\partial \sigma_{p,t+1}}{\partial \mu_{t+1}} \frac{\partial V (x_{t+1})}{\partial \sigma_{p,t+1}} + \frac{\partial \sigma_{x,t+1}}{\partial \mu_{t+1}} \frac{\partial V (x_{t+1})}{\partial \sigma_{x,t+1}} \right) \right] +
\]

\[ E_t \left[ \beta_t V (x_{t+1}) \frac{\partial f (s_{t+1} | x_t, u_t)}{\partial s_{t+1}} \right] \]

First, consider the term \( \frac{\partial V (x_{t+1})}{\partial \mu_{t+1}} \). By the envelope theorem one may write

\[ \frac{\partial V (x_{t+1})}{\partial \mu_{t+1}} = \frac{\partial \{ F (c (x_{t+1})) + E_{t+1} [\beta_{t+1} V (x_{t+2})] \}}{\partial \mu_{t+1}} \]

\[ = \frac{\partial E_{t+1} [\beta_{t+1} V (x_{t+2})]}{\partial \mu_{t+1}} \]

\[ = \beta_{t+1} \left( E_{t+1} \left[ \frac{\partial f_{t+1} (a_{t+2}, s_{t+1})}{\partial \mu_{t+1}} V (x_{t+2}) + \frac{\partial V (x_{t+2})}{\partial x_{t+2}} \frac{\partial x_{t+2}}{\partial \mu_{t+1}} \right] \right) \]

\[ = \beta_{t+1} E_{t+1} \left[ \frac{\partial f_{t+1} (a_{t+2}, s_{t+1})}{\partial \mu_{t+1}} V (x_{t+2}) \right] \]

It may be verified that

\[ \frac{\partial f_{t+1} (a_{t+2}, s_{t+1})}{\partial \mu_{t+1}} = \frac{a_{t+2} - E_{t+1} [a_{t+2}]}{Var_{t+1} [a_{t+2}]} \]

with

\[ E_{t+1} [a_{t+2}] = \omega_{a,t+1} + \hat{\mu}_{t+1} \]

\[ Var_{t+1} [a_{t+2}] = \sigma_{t+1}^2 + \sigma_v^2 \]

Given \( V (x_{t+2}) \) satisfies condition \( (1) \) one may write

\[ \frac{\partial V (x_{t+1})}{\partial \mu_{t+1}} = \beta_{t+1} Cov_{t+1} \left[ \frac{a_{t+2} - E_{t+1} [a_{t+2}]}{Var_{t+1} [a_{t+2}]}, V (x_{t+2}) \right] \]

\[ = \beta_{t+1} \frac{Cov_{t+1} [a_{t+2}, V (x_{t+2})]}{Var_{t+1} [a_{t+2}]} \]

\[ = \beta_{t+1} \frac{Var_{t+1} [a_{t+2}]}{Var_{t+1} [a_{t+2}]} E_{t+1} \left[ \frac{\partial V (x_{t+2})}{\partial x_{t+2}} \frac{\partial x_{t+2}}{\partial a_{t+2}} \right] \]

\[ = \beta_{t+1} E_{t+1} \left[ \frac{\partial V (x_{t+2})}{\partial y_{t+2}} \right] \]

Further, using the envelope theorem

\[ \frac{\partial V (x_{t+2})}{\partial y_{t+2}} = F' (c_{t+2}) \cdot Y_{t+2} \]
we obtain

\[
\frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} = \beta_{t+1}E_{t+1} [F'(c_{t+2}) \cdot Y_{t+2}]
\]

\[
= F'(c_{t+1}) G'(k_{t+1}) E_{t+1} \left[ A_{t+2}G'(k_{t+1}) \frac{\beta_{t+1} F'(c_{t+1})}{G'(k_{t+1})} \right]
\]

\[
\frac{\partial V(x_{t+1})}{\partial \mu_{t+1}} = F'(c_{t+1}) G'(k_{t+1}) \frac{\beta_{t+1} F'(c_{t+1})}{G'(k_{t+1})}
\]

Next, consider the term \( \frac{\partial V(x_{t+1})}{\partial \sigma^2_{t+1}} \). By the envelope theorem we may write

\[
\frac{\partial V(x_{t+1})}{\partial \sigma^2_{t+1}} = \frac{\partial F(c(x_{t+1}))}{\partial \sigma^2_{t+1}} + E_{t+1} [\beta_{t+1} V(x_{t+2})]
\]

\[
= \frac{\partial E_{t+1} [\beta_{t+1} V(x_{t+2})]}{\partial \sigma^2_{t+1}}
\]

\[
= \beta_{t+1} E_{t+1} \left[ \frac{\partial f_{t+1}(a_{t+2})}{\partial \sigma^2_{t+1}} \right] V(x_{t+2}) + \beta_{t+1} E_{t+1} \left[ \frac{\partial V(x_{t+2})}{\partial \sigma^2_{t+1}} \right] \frac{\partial x_{t+2}}{\partial \sigma^2_{t+1}}
\]

It may be verified that

\[
\frac{\partial f_{t+1}(a_{t+2})}{\partial \sigma^2_{t+1}} = \frac{1}{2} \left( (a_{t+2} - E_{t+1} [a_{t+2}])^2 - \text{Var}_{t+1} [a_{t+2}] \right)
\]

Given \( V(x_{t+2}) \) satisfies condition [1] we obtain

\[
\frac{\partial V(x_{t+1})}{\partial \sigma^2_{t+1}} = \frac{1}{2} \left( (a_{t+2} - E_{t+1} [a_{t+2}])^2 - \text{Var}_{t+1} [a_{t+2}] \right) \frac{\partial f_{t+1}(a_{t+2})}{\partial \sigma^2_{t+1}} V(x_{t+2}) + \left( \text{Cov}_{t+1} \left[ a_{t+2}, V(x_{t+2}) \right] \right) \frac{\partial V(x_{t+2})}{\partial \sigma^2_{t+1}} \frac{\partial x_{t+2}}{\partial \sigma^2_{t+1}}.
\]
Hence, we obtain
\[
\frac{\partial V(x_{t+1})}{\partial h_{t+1}} = \frac{\partial F(c(x_{t+1})) + E_{t+1}[\beta_{t+1}V(x_{t+2})]}{\partial h_{t+1}} = -F'(c_{t+1}) n_h (t_{t+1}, h_{t+1}) + \frac{\partial E_{t+1}[\beta_{t+1}V(x_{t+2})]}{\partial h_{t+1}}
\]
\[
= -F'(c_{t+1}) n_h (t_{t+1}, h_{t+1}) + E_{t+1} \left[ \beta_{t+1} \frac{\partial V(x_{t+2})}{\partial x_{t+2}} \frac{\partial x_{t+2}}{\partial h_{t+1}} \right]
\]
\[
= F'(c_{t+1}) n_e (t_{t+1}, h_{t+1})
\]

since, given the functional form assumed, \( n_e (t_e, h_e) = -n_h (t_e, h_e) \). Finally consider the term \( E_t \left[ V(x_{t+1}) \frac{\partial f_i(s_{t+1})}{\partial s_{t+1}} / f_i (s_{t+1}) \right] \). It may be verified that
\[
\frac{\partial f_i(s_{t+1})}{\partial s_{t+1}} = -\frac{\sigma^2_{\nu, t}}{2} \frac{s_{t+1}^2 - \text{Var}_{s_{t+1}}[s_{t+1}]}{(\sigma^2_{\mu, t} + \sigma^2_{\nu, t})^2}
\]

Hence, we obtain
\[
E_t \left[ V(x_{t+1}) \frac{\partial f_i(s_{t+1})}{\partial s_{t+1}} / f_i (s_{t+1}) \right] = -\frac{\sigma^2_{\nu, t}}{2} \left( E_t \left[ s_{t+1}^2 V(x_{t+1}) \right] - \text{Var}_{s_{t+1}}[s_{t+1}] E_t[V(x_{t+1})] \right)
\]
\[
= -\frac{\sigma^2_{\nu, t}}{2} \left( E_t \left[ s_{t+1}^2 V(x_{t+1}) \right] - \text{Var}_{s_{t+1}}[s_{t+1}] E_t[V(x_{t+1})] \right)
\]
\[
= -\frac{\sigma^2_{\nu, t}}{2} \left( \text{Cov}_{s_{t+1}}[V(x_{t+1}) s_{t+1}] - \text{Var}_{s_{t+1}}[s_{t+1}] E_t[V(x_{t+1})] \right)
\]
\[
= -\frac{\sigma^2_{\nu, t}}{2} \left( \text{Var}_{s_{t+1}}[s_{t+1}] E_t \left[ \frac{\partial V(x_{t+1})}{\partial s_{t+1}} \right] \frac{s_{t+1}}{s_{t+1}^2 + \sigma^2_{\nu, t}} \right)
\]
\[
= -\frac{1}{2} E_t \left[ \frac{\partial V(x_{t+1})}{\partial s_{t+1}} \frac{s_{t+1}^2 + \sigma^2_{\nu, t}}{\left( 1 + \sigma^2_{\nu, t} \right)^2} \right]
\]
Note that
\[
\frac{\partial \hat{\mu}_{t+1}}{\partial t} = \frac{s_{t+1} \sigma_{\mu,t}^2}{(1 + t \sigma_{\mu,t}^2)^2}
\]
Hence, we may write
\[
E_t \left[ V(x_{t+1}) \frac{\partial f_t(s_{t+1})}{\partial x_t} \right] = -\frac{1}{2} E_t \left[ \frac{\partial V(x_{t+1}) \partial \hat{\mu}_{t+1}}{\partial t} \right].
\]
Substituting these results into the FONC yields
\[
1 = E_t \left[ m_{t+1} \left( \frac{\partial \hat{\mu}_{t+1}}{\partial t} \frac{\partial f_t(s_{t+1})}{\partial x_t} + \frac{\partial \sigma_{t+1}^2}{\partial t} \frac{Cov_{t+1}[a_{t+2}, m_{t+2} Y_{t+2}]}{Var_{t+1}[a_{t+2}]} \right) + \delta h n_t (t_{t+1}, h_{t+1}) \right] / n_t (t, h_t)
\]
where \( m_{t+1} \equiv \frac{\beta_t F'(c_{t+1})}{F'(c_t)} \)
and
\[
\frac{\partial \hat{\mu}_{t+1}}{\partial t} = \frac{s_{t+1} \sigma_{\mu,t}^2}{(1 + t \sigma_{\mu,t}^2)^2}
\]
\[
\frac{\partial \sigma_{t+1}^2}{\partial t} = -\sigma_{\mu,t}^2 \frac{\sigma_{\mu,t}^2}{(1 + t \sigma_{\mu,t}^2)^2}
\]
In terms of the notation previously used we obtain
\[
E_t \left[ m_{t+1} \frac{\Pi_{t+1} + \delta_{t+1} MC_{t+1}}{MC_t} \right] = 1,
\]
with
\[
\Pi_{t+1} = \frac{1}{2} \left( \frac{s_{t+1} G(k_{t+1})}{\sigma_{\mu,t}^2 G'(k_{t+1})} - \frac{Cov_{t+1}[a_{t+2}, m_{t+2} Y_{t+2}]}{Var_{t+1}[a_{t+2}]} \right)
\]
\[
MC_t = -n_t (t, h_t) \left/ \frac{\partial \sigma_{t+1}^2}{\partial t} \right.
\]
\[
\delta_{t+1} = \delta_h \left/ \frac{\partial \sigma_{t+1}^2}{\partial t} \right.
\]
\[
\frac{\partial \sigma_{t+1}^2}{\partial t} = -\sigma_{\mu,t}^2 \frac{\sigma_{\mu,t}^2}{(1 + t \sigma_{\mu,t}^2)^2}
\]

and where $\hat{\sigma}_{t+1}^2 = \hat{\sigma}_{t+1}^2 (\hat{s}_{\mu,t}, t)$. Under linear utility with exogenous, $F_t$ measurable discount factors $m_{t+1} = \beta_t$,

$$Cov_{t+1} [a_{t+2}, m_{t+2} Y_{t+2}] = Var_{t+1} [a_{t+2}] E_{t+1} [m_{t+2} A_{t+2}] G (k_{t+1})$$

The FONC for capital,

$$E_{t+1} [m_{t+2} A_{t+2}] G' (k_{t+1}) = 1,$$

allows us to write

$$Cov_{t+1} [a_{t+2}, m_{t+2} Y_{t+2}] = \frac{G (k_{t+1})}{G' (k_{t+1})} Var_{t+1} [a_{t+2}] .$$

Substituting this result into the previous equation yields

$$\Pi_{t+1}^{\gamma=0} = \frac{1}{2} \frac{G (k_{t+1})}{G' (k_{t+1})} \left( \frac{s_{t+1}}{\sigma_{\mu,t}^2} - 1 \right)$$

$$= \frac{1}{2} E_{t+1} [m_{t+2} A_{t+2}] G (k_{t+1}) \left( \frac{s_{t+1}}{\sigma_{\mu,t}^2} - 1 \right)$$

where $k_{t+1}$ is given by

$$k_{t+1} = G^{-1} \left( \frac{1}{E_{t+1} [m_{t+2} A_{t+2}]} \right)$$

with $G^{-1}$ denoting the inverse of $G'$. For $G (k_t) = k^{1-\eta}$, with $0 < \eta < 1$ one obtains

$$k_{t+1} = (E_{t+1} [m_{t+2} A_{t+2}] (1-\eta))^{1/\eta}$$

$$G (k_{t+1}) = (E_{t+1} [m_{t+2} A_{t+2}] (1-\eta))^{1-\eta/\eta}$$

Substituting yields

$$\Pi_{t+1}^{\gamma=0} = \frac{1}{2} (1-\eta)^{1-\eta/\eta} (E_{t+1} [m_{t+2} A_{t+2}])^{1/\eta} \left( \frac{s_{t+1}}{\sigma_{\mu,t}^2} - 1 \right)$$

Taking conditional expectations

$$E_t \left[ \Pi_{t+1}^{\gamma=0} \right] = \frac{1}{2} (1-\eta)^{1-\eta/\eta} E_t \left[ (E_{t+1} [m_{t+2} A_{t+2}]^{1/\eta} \left( \frac{s_{t+1}}{\sigma_{\mu,t}^2} - 1 \right) \right]$$

where

$$E_t \left[ (m_{t+2} E_{t+1} [A_{t+2}]^{1/\eta} \left( \frac{s_{t+1}}{\sigma_{\mu,t}^2} - 1 \right) \right]$$

$$Cov_t \left[ m_{t+2}^{1/\eta} e^{\frac{1}{2} \{ \omega_{n,t_i+1} + \bar{\mu}_{t+1} + \frac{1}{2} (\hat{s}_{\mu,t_i+1}^2 + \hat{s}_{\sigma_i,t_i+1}^2) \}, s_{t+1} \right]$$

$$= Var [s] E_t \left[ m_{t+2}^{1/\eta} e^{\frac{1}{2} \{ \omega_{n,t_i+1} + \bar{\mu}_{t+1} + \frac{1}{2} (\hat{s}_{\mu,t_i+1}^2 + \hat{s}_{\sigma_i,t_i+1}^2) \}} \right]$$
Conditinal on $\text{Cov} [s_{t+1}, m_{t+2}] = 0$, such that $m_{t+2} = \beta_{t+1}$ does not reveal any
information about the hidden productivity state, we may write

$$E_t \left[ (m_{t+2} E_{t+1} [A_{t+2}])^{\frac{1}{\eta}} \left( \frac{s_{t+1}}{\sigma_{\mu,t}} - 1 \right) \right]$$

$$= E_t \left[ m_{t+2}^{\frac{1}{\eta}} e^{\frac{1}{\eta} \left( \omega_{a,t+1} + \frac{\partial \mu_{t+1}}{\partial s_{t+1}} s_{t+1} + \frac{1}{2} \left( \hat{\sigma}_{e+1}^2 + \sigma_{v,t+1}^2 \right) \right)} \right]$$

$$= \frac{1}{\sigma_{\mu,t}} \text{Cov}_t \left[ m_{t+2}^{\frac{1}{\eta}} e^{\frac{1}{\eta} \left( \omega_{a,t+1} + \frac{\partial \mu_{t+1}}{\partial s_{t+1}} s_{t+1} + \frac{1}{2} \left( \hat{\sigma}_{e+1}^2 + \sigma_{v,t+1}^2 \right) \right)}, s_{t+1} \right] -$$

$$E_t \left[ m_{t+2}^{\frac{1}{\eta}} e^{\frac{1}{\eta} \left( \omega_{a,t+1} + \frac{\partial \mu_{t+1}}{\partial s_{t+1}} s_{t+1} + \frac{1}{2} \left( \hat{\sigma}_{e+1}^2 + \sigma_{v,t+1}^2 \right) \right)} \right]$$

$$= E_t \left[ \left( E_t [m_{t+2} A_{t+2}] \right)^{\frac{1}{\eta}} \right] \frac{1}{\eta} (1 - \eta)$$

such that

$$E_t \left[ \Pi_{t+1}^{\gamma=0} \right] = E_t \left[ \frac{((1 - \eta) m_{t+2} E_t [A_{t+2}])^{\frac{1}{\eta}}}{2\eta} \right]$$

which is always greater than zero since since $0 < \eta < 1$. Further, given that

$$m_{t+2}$$

is $\mathcal{F}_t$ measurable and given that $\sigma_{\epsilon,t+1}^2 = \sigma_{\epsilon}^2$ we obtain

$$E_t \left[ \Pi_{t+1}^{\gamma=0} \right] = \frac{((1 - \eta) m_{t+2})^{\frac{1}{\eta}} e^{\frac{1}{\eta} \left( \omega + \phi \omega_{a,t} + \frac{1}{2} \left( \hat{\sigma}_{e+1}^2 + \sigma_{v,t+1}^2 \right) + \eta \left( \sigma_{\epsilon,t+1}^2 / \text{Var}_t [s_{t+1}] \right) \right)}{2\eta}$$

$\square$