

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Uncertain Growth Cycles, Corporate Investment, and Dynamic Hedging

Permalink

<https://escholarship.org/uc/item/38p0r10v>

Author

Yonce, Adam

Publication Date

2010

Peer reviewed|Thesis/dissertation

Uncertain Growth Cycles, Corporate Investment, and Dynamic Hedging

by

Adam Theodore Yonce

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Business Administration

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Richard Stanton, Chair
Professor Roger Craine
Professor Johan Walden

Spring 2010

Uncertain Growth Cycles, Corporate Investment, and Dynamic Hedging

Copyright 2010
by
Adam Theodore Yonce

Abstract

Uncertain Growth Cycles, Corporate Investment, and Dynamic Hedging

by

Adam Theodore Yonce

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Richard Stanton, Chair

In the theory of finance, uncertainty plays a crucial role. Economists often use the terms uncertainty and volatility interchangeably, yet volatility is not the only form of uncertainty. Firms face uncertainty about whether the economy is in an expansionary or recessionary state, industries face regulatory uncertainty, and individuals face uncertainty about risk premia. In this dissertation, I consider the role that uncertainty about growth rates, regulatory policy, and risk premia play in the investment decisions of firms and individuals. The key theme linking the three chapters is learning in dynamic environments.

In Chapter 1, I study the effects of demand growth uncertainty on corporate investment decisions. In particular, how does uncertainty about the state of the economy and the state of demand growth affect a firm's decision to allocate capital to irreversible investment projects? In the model, firms are able to choose both the timing and scale of their investments, and the optimal scale will depend on the unobserved state of demand growth. This second decision gives rise to an incentive to delay investment that does not exist in standard real option models: When investment is irreversible, firms risk allocating a sub-optimal level of capital to a project. Theoretically, I show how this incentive to delay is closely linked to the benefits of learning about the economy. Empirically, using estimated probabilities filtered from GDP growth, I find that 1) beliefs about the economy inform corporate investment decisions, and 2) the relationship between investment and beliefs is quadratic.

In Chapter 2, I study an empirical extension of the model. Many industries in the United States face regulatory uncertainty, and a natural conjecture is that increased regulatory uncertainty has a dampening effect on investment if 1) regulatory policy affects the cash flows of the firm, 2) firms have flexibility over the scale of their investments, and 3) regulatory uncertainty resolves quickly. While regulatory uncertainty is not observable, I consider two proxies: A variable indicating Presidential election years, and a variable indicating divided government. The former is meant to capture policy uncertainty associated with the possibility of a change in government, while the latter is meant to capture policy uncertainty associated with ideological variance.

Empirically, both measures are associated with a decrease in corporate investment rates, consistent with the theoretical framework. The second purpose of this chapter is to highlight the dangers of making inferences about investment using inconsistent estimators and regressions that fail to account for plausible alternative hypotheses. Previous work linking investment to the political cycle relies on least squares estimators that are inconsistent because the firm-specific control variables are endogenous to the investment decision. For a specific sub-sample of non-manufacturing firms, I show that least squares estimates easily reject the null hypothesis, while consistent first-difference estimates fail to do so. Finally, I include a control for the fiscal environment of the federal government, which helps to uncover important dynamics between investment, the budget deficit, and the election cycle.

In chapter 3, I consider the currency hedging problem of a risk-averse international investor who faces an unobservable currency risk premium. A non-zero risk premium introduces a speculative motive for holding foreign currency in the optimal portfolio, and a time-varying risk premium introduces a market-timing strategy. Uncertainty about the stochastic properties of the risk premium significantly tames both the speculative and market timing components, especially at long investment horizons, and the optimal hedge approaches a complete hedge as risk aversion and the investment horizon increase. However, an investor who ignores the risk premium and fully hedges foreign investments faces a substantial opportunity cost because she forgoes the benefits of dynamic learning.

Contents

List of Figures	iii
List of Tables	iv
1 A Model of Investment Under Demand Growth Uncertainty	1
1.1 Introduction	1
1.2 Theoretical Framework	3
1.2.1 Market Dynamics	3
1.2.2 The Firm's Learning Problem	6
1.2.3 Project Valuation	11
1.3 Numerical Analysis	15
1.3.1 Calibration to US GDP Growth	15
1.4 Empirical Analysis	18
1.4.1 Parameter Estimation	19
1.4.2 Empirical Analysis of Corporate Investment	21
1.5 Concluding Remarks	26
1.6 Figures and Tables	28
2 An Empirical Analysis of Investment Over the Political Cycle	43
2.1 Introduction	43
2.2 Theoretical Review	47
2.3 Empirical Framework	50
2.3.1 Empirical Specification	50
2.3.2 Data Description	53
2.3.3 Estimation	54
2.4 Concluding Remarks	57
2.5 Tables	59
3 Dynamic Currency Hedging	68
3.1 Introduction	68
3.2 Currency Hedging with Constant Risk Premia	70
3.2.1 Asset Market Dynamics	70

3.2.2	The Investor's Learning Problem	71
3.2.3	The Investor's Asset Allocation Problem	73
3.2.4	Value Function Solution	75
3.2.5	Analysis	76
3.3	A General Model of Portfolio Choice	77
3.3.1	Asset Market Dynamics	77
3.3.2	The Inference Problem	79
3.3.3	The Portfolio Problem	81
3.3.4	Value Function	82
3.4	Numerical Analysis	83
3.5	The Cost of Suboptimal Hedging	84
3.5.1	Expected Utility of Terminal Wealth	84
3.5.2	The Cost of Complete Hedging	86
3.6	Concluding Remarks	87
3.7	Figures	88
A	Appendix to Chapter 1	97
A.1	The Filtering Equations	97
A.2	Properties of the Learning Process	102
A.2.1	Conditional Moments	102
A.3	Bellman Equation	103
A.4	Maximum Likelihood Estimation	104
A.5	Numerical Solution to the PDE	106
B	Appendix to Chapter 3	108
B.1	ODE for the Posterior Variance	108
B.2	Value Function in the General Model	109

List of Figures

1.1	Steady-state probability density for π	28
1.2	Probability that the economy is in a recessionary state	29
1.3	Empirical and theoretical densities of posterior beliefs	30
1.4	Free boundary solution to the investment model	31
1.5	The effect of volatility	32
1.6	The effect of signal quality	33
1.7	The effect of shift intensity	34
1.8	Expected investment conditional on beliefs about the economy	35
3.1	Optimal currency hedge ratios, Simple Model	88
3.2	Posterior variance as a function of time	89
3.3	Optimal hedge ratio as a function of the risk premium	90
3.4	Optimal currency hedge ratios, General Model	91
3.5	Opportunity cost of complete hedging	92

List of Tables

1.1	Description of COMPUSTAT variables	36
1.2	Regime-Switching Parameters for GDP Growth	37
1.3	Summary Statistics - Investment, Cash Flow, and Tobin's Q	38
1.4	IV Regressions - Capital Expenditures and Annual Beliefs	39
1.5	IV Regressions - Capital Expenditures and Average Quarterly Beliefs	40
1.6	IV Regressions - Capital Expenditures and Acquisitions, Annual Beliefs	41
1.7	IV Regressions - Capital Expenditures and Acquisitions, Quarterly Beliefs	42
2.1	Description of FRED variables	59
2.2	Summary Statistics - Investment, Cash Flow, and Tobin's Q	60
2.3	OLS Regressions - Investment and Political Cycles	61
2.4	Least Squares Regressions	62
2.5	Least Squares vs. Instrumental Variables	63
2.6	Capital Expenditures, All Firms	64
2.7	Capital Expenditures and Acquisitions, All Firms	65
2.8	Capital Expenditures, Manufacturing Firms	66
2.9	Capital Expenditures and Acquisitions, Manufacturing Firms	67

Acknowledgments

For helpful comments and discussions, I would like to thank Roger Craine, Ernesto Dal Bo, Sara Holland, Rich Lyons, Richard Stanton, and Johan Walden.

Additional thanks go to Kim Guilfoyle, Lagunitas Brewing Company, and the California Bluegrass Association.

Finally, I would like to thank my parents, John and Marion Yonce, for encouraging me to pursue a Ph.D.

Chapter 1

A Model of Investment Under Demand Growth Uncertainty

1.1 Introduction

In the theory of real options, uncertainty plays a crucial role. Investing in a project is equivalent to exercising a call option; as in equity options, increased volatility is associated with increased potential upside, but without any increased downside. All else equal, higher volatility leads to higher exercise thresholds and greater project values.

Yet volatility is not the only form of uncertainty. Firms also face uncertainty about whether the economy is in an expansionary or recessionary state, and certain industries also face regulatory uncertainty. Indeed, in Chapter 2 I provide evidence that corporate cash flows differ systematically across policy environments, and that firms systematically decrease investment rates during periods of higher policy uncertainty.

This chapter is a study the effects of demand growth uncertainty on corporate investment decisions. In particular, how does uncertainty about the state of the economy and the state of demand growth affect a firm's decision to allocate capital to irreversible investment projects? Firms may find it optimal to forgo NPV-positive projects while retaining their right to invest at a later date, if delaying allows them to learn about the demand.

I study the link between economic uncertainty and investment through a dynamic real option model with two key ingredients. First, firms do not have perfect information about the economic environment, which governs the growth rate of demand and cash flows, and must form a belief about whether the economy is expanding or contracting. Beliefs about the economy then inform the firm's valuation of a perpetual stream of cash flows.

Second, unlike standard real option models, firms are able to choose the scale of their investments. In addition to choosing the optimal time to invest, firms must

choose the optimal amount of capital to invest. When beliefs are optimistic, so that estimated growth rates and present values are high, firms will want to commit a large amount of capital to take advantage of high growth. When beliefs are pessimistic, however, firms commit less capital since valuations are lower and projects appear less attractive.

Under this setup, the model generates two key predictions:

1. Firms will invest when they have precise beliefs about the economic environment, regardless of whether beliefs are optimistic or pessimistic, but will delay when uncertainty is highest.
2. The incentive to delay is strongest when uncertainty resolves quickly.

The first prediction is a consequence of a firm's ability to choose the scale of its projects. The amount of capital that maximizes NPV differs in the two economic environments. When investment is irreversible, firms must guard against allocating an inefficient amount of capital to a project. Therefore, they may optimally delay projects to gather more information about demand. This result is intuitive, but it depends critically on allowing the scale of the project to vary. In standard real option models, where firms face just a timing problem, firms will only invest if uncertainty resolves in the favorable direction. When scale is a choice variable, firms invest even when uncertainty resolves unfavorably; they just choose smaller projects.

The second prediction is a consequence of learning. Forgoing an NPV-positive project can be optimal if delaying investment allows the firm to learn about the economic environment. The benefit can be substantial if delaying allows the firm to make a more informed capital allocation choice. When uncertainty resolves slowly, however, there is little to learn. In this case, the opportunity cost of delaying investment will outweigh the benefits gained from learning.

I test the predictions on U.S. firms by fitting a switching model to the growth of real gross domestic product and estimating a time series of associated probabilities that the economy is in a high-growth state. Together with the estimated probabilities, the parameters of the switching model provide evidence of a bi-modal distribution in beliefs about the economy, and quick resolution of uncertainty. Consistent with the model, I find that 1) beliefs about the economy do indeed inform corporate investment decisions, and 2) the relationship between investment and beliefs is quadratic. The intuition is that the relationship between uncertainty and beliefs is also quadratic - if the probability that the economy is in a high-growth state is given by π , then uncertainty is given by $\pi(1 - \pi)$. Therefore, as information about the economy becomes less precise, firms find it optimal to delay investment in order to learn about the economy. However, this incentive to delay is substantially diminished if learning is slow.

The theoretical framework builds on previous work by Bernanke (1983) and Klein (2007). Bernanke introduced Bayesian analysis into investment decisions in order to

explain aggregate investment cycles. In his model, the economic policy environment directly affects the relative value of mutually exclusive investments, and investors form beliefs about how long a policy persists. A key feature of his analysis is that beliefs evolve monotonically and deterministically, so there is a sectoral shift in investment and cyclical behavior. In his example of an energy cartel, where a firm chooses between energy-intensive and energy-saving investments, there is a period of no investment before the shift as beliefs about cartel persistence evolve from low probability to high probability.

Unlike Bernanke (1983), in this paper beliefs evolve stochastically, and potentially non-monotonically, in response to news about the economy. Because the economic environment can always change, there is a tendency for beliefs to mean-revert, and uncertainty never disappears completely.¹

Klein (2007) studies a model where the total value of the investment project is the stochastic variable. In his model, firms have incomplete information over the drift of the value process - they know it can take one of only two values, but they are uncertain which value is the true growth rate. Because there are no regime shifts in his model, firms eventually completely uncover the true stochastic process. In this paper, the primitive stochastic variable is demand, and the growth rate is allowed to randomly switch between high-growth and low-growth states. Because the growth rate can shift between two regimes, firms are never able to completely learn the true stochastic process, although they are able to form very precise beliefs. An additional feature of the setup is that the model endogenously generates stochastic volatility in project values.

I now turn to a discussion of the theoretical framework and the solution to the firm's investment problem.

1.2 Theoretical Framework

1.2.1 Market Dynamics

A firm has the option to invest K units of capital in a project. Investment is irreversible, and I assume that capital does not depreciate. After investment, the firm produces $F(K) = \log(1 + K)$ units of output in perpetuity. The output price is P , so that per-period cash flows are $P(t)F(K)$.

The firm is risk-neutral, and its objective is to choose an investment time τ , as well as an investment level K , to maximize the expected net present value (NPV) of

¹Prediction markets provide a useful example of beliefs evolving in a stochastic and non-monotonic way. See Wolfers and Zitzewitz (2006) for a discussion of when we can interpret prediction market prices as probabilities.

cash flows,

$$G = \max_{K, \tau} E_0 \left[\int_{\tau}^{\infty} e^{-ru} P(u) F(K_{\tau}) du - e^{-r\tau} p_K K_{\tau} \right]. \quad (1.1)$$

The integral term in (1.1) is the expected present value of future cash flows; p_K is the per-unit price of capital, so that $e^{-r\tau} p_K K_{\tau}$ is the discounted cost of the investment. As with standard real option models, the expected NPV of delaying investment may exceed the expected NPV of investing immediately. In this case, firms find it optimal to forgo positive NPV projects while retaining their right to invest at a future date.

For simplicity, I assume the price of capital is non-random and normalize p_K to unity.² The price of the output good is stochastic and is governed by a downward-sloping inverse demand curve

$$P(t) = Y(t)F(K)^{-1/\gamma} \quad (1.2)$$

where γ is the price elasticity of demand.³ Y is a stochastic demand shock and follows a geometric process,

$$\frac{dY}{Y} = \mu(\theta)dt + \sigma dZ_Y \quad (1.3)$$

where Z_Y is a standard Brownian motion and $\theta(t) \in \{\theta_h, \theta_l\}$ is a state variable describing the economic environment, corresponding to ‘high growth’ and ‘low growth.’ This is meant to capture periods of economic expansion and recession. θ , itself a random process, follows a two-state Markov process with transition matrix

$$\mathbf{\Lambda} = \begin{pmatrix} -\lambda_{lh} & \lambda_{lh} \\ \lambda_{hl} & -\lambda_{hl} \end{pmatrix} \quad (1.4)$$

where λ_{lh} is the hazard rate for transition from the low growth state to the high growth state, and λ_{hl} is the hazard rate for transition from the high growth state to the low growth state.

As a consequence of the two-state Markov assumption, θ has dynamics

$$d\theta = (\theta_h - \theta_l) \left[1 - \frac{2(\theta(t) - \theta_h)}{\theta_l - \theta_h} \right] dq(\theta)$$

where dq is a Poisson process governed by transition matrix $\mathbf{\Lambda}$. Under this process, θ jumps between states at random, discrete intervals, and the time between jumps follows an exponential distribution. The arrival rate of the Poisson process (also the

²Adapting the model to include stochastic p_K is conceptually straightforward. See Chapter 6 of Dixit and Pindyck (1994) for the standard method of transforming a two-variable investment problem to a single-variable problem.

³The price elasticity of demand ϵ is defined as $\epsilon = \frac{dQ}{dP} \frac{P}{Q}$. For an iso-elastic demand function of the form $Q = \left(\frac{P}{Y}\right)^{-\gamma}$, this simplifies to $\epsilon = -\gamma$.

rate parameter of the exponential distribution) is $\lambda(\theta)$, with

$$\lambda(\theta) = \begin{cases} \lambda_{lh} & \text{if } \theta = \theta_l \\ \lambda_{hl} & \text{if } \theta = \theta_h \end{cases}, \quad (1.5)$$

so that the mean time between jumps is $1/\lambda(\theta)$.

In this model, there is nothing special about the numerical values of θ ; the only requirement is that $\theta_h \neq \theta_l$. In the numerical analysis of the model it will be convenient to normalize to $\theta_l = 0$ and $\theta_h = 1$. In this case, the process for θ simplifies to

$$d\theta = [1 - 2\theta(t)] dq(\theta).$$

I make two key assumptions about Y and θ :

1. Demand Y is observable, but the Brownian motion Z_Y is unobservable.
2. θ follows a hidden (i.e. unobservable) Markov process. The parameters of the process $(\theta_l, \theta_h, \mathbf{\Lambda})$ are known, but the current state $\theta(t)$ is unknown.

Given these assumptions, the firm cannot directly observe the economic environment; that is, they have incomplete information. While Y is observable, growth (or decay) in Y due to θ is indistinguishable from growth (or decay) due to Z_Y . To gain an intuitive understanding, briefly suppose that $\mu(\theta_h) > 0$ and $\mu(\theta_l) < 0$. Then growth in Y could be due to a strong overall economy ($\theta(t) = \theta_h$), a positive demand shock ($dZ_Y > 0$), or both. Likewise, decay in Y could be due to a weak overall economy ($\theta(t) = \theta_l$), a negative demand shock ($dZ_Y < 0$), or both. The important feature is that a firm observing the path of $Y(t)$ would not be able to perfectly distinguish between the alternatives.

However, while not having complete information, the firm may see a noisy signal s about the economy,

$$ds = \theta dt + \eta \rho dZ_Y + \eta \sqrt{1 - \rho^2} dZ_s, \quad (1.6)$$

or, intuitively, the signal of θ is equal to the true value plus some random noise. The diffusion parameter η captures noise, and Z_s is a standard Brownian motion orthogonal to Z_Y . The parameter ρ captures any possible correlation between demand and signals. Indeed, positive demand shocks are likely to be correlated with positive signals about θ . Examples of noisy signals include economic indicators such as equity market returns and employment data, and may also include actions and statements by monetary authorities, regulators, and politicians.

1.2.2 The Firm's Learning Problem

The Filtering Equations

While the firm does not directly observe the economic environment, they may use observations of Y and s to form beliefs about θ . In particular, beliefs are summarized by the posterior probability π ,

$$\pi(t) = \text{Prob}(\theta = \theta_h | \mathcal{F}_t), \quad (1.7)$$

i.e. the probability that the economic growth environment is high, conditional on the current information set. The firm updates beliefs via Bayes's Rule as new observations of demand and signals arrive. Armed with posterior beliefs, the firm then forms expectations about the current economic environment and the growth rate of demand:

$$\begin{aligned} \hat{\theta} &= \pi\theta_h + (1 - \pi)\theta_l \\ \hat{\mu} &= \pi\mu_h + (1 - \pi)\mu_l \end{aligned}$$

where, to simplify notation, I define $\mu_h = \mu(\theta_h)$ and $\mu_l = \mu(\theta_l)$.

Optimal filtering equations (i.e. the stochastic processes for posterior beliefs π and expected growth rates $\hat{\mu}$) follow from, in the case of one signal, Theorem 9.1 of Liptser and Shiryaev (2001). Lemma 1 of Veronesi (2000) generalizes to a multivariate setting. Equation (A.1.13) in Appendix A gives the stochastic process for π when the firm observes both demand Y and the noisy signal s :

$$\begin{aligned} d\pi &= \kappa(\bar{\pi} - \pi)dt \\ &+ \pi(1 - \pi) \left(\frac{\mu_h - \mu_l}{\sigma^2(1 - \rho^2)} - \frac{\theta_h - \theta_l}{\eta\sigma} \frac{\rho}{1 - \rho^2} \right) \left(\frac{dY}{Y} - \hat{\mu}dt \right) \\ &+ \pi(1 - \pi) \left(\frac{\theta_h - \theta_l}{\eta^2(1 - \rho^2)} - \frac{\mu_h - \mu_l}{\eta\sigma} \frac{\rho}{1 - \rho^2} \right) (ds - \hat{\theta}dt) \end{aligned} \quad (1.9)$$

with

$$\begin{aligned} \kappa &= \lambda_{lh} + \lambda_{hl} \\ \bar{\pi} &= \frac{\lambda_{lh}}{\lambda_{lh} + \lambda_{hl}}. \end{aligned}$$

The firm optimally updates its beliefs about the economic environment according to equation (1.9), which contains two components: A deterministic, mean-reverting component, which pulls beliefs towards a long-run value of $\bar{\pi}$ at rate κ , and a stochastic component driven by unexpected shocks to dY and ds . Intuitively, the stochastic component represents news from demand (or price) and signals. When dY and ds are

large relative to expectations (in absolute value), signals convey a lot of information about θ , and the revision in beliefs is potentially large.

When the firm only observes demand, the filtering equation simplifies to

$$d\pi = \kappa(\bar{\pi} - \pi)dt + \pi(1 - \pi)\frac{\mu_h - \mu_l}{\sigma^2} \left[\frac{dY}{Y} - \hat{\mu}dt \right].$$

While equation (1.9) fully characterizes the learning process of the firm, the stochastic equations for prices and signals are still written in terms of unobservable parameters and shocks, specifically $\theta(t)$, $dZ_Y(t)$, and $dZ_s(t)$. However, consider the transformations defined in equation (A.1.15):

$$d\hat{Z}_Y = \frac{\mu(\theta) - \hat{\mu}}{\sigma}dt + dZ_Y \quad (1.11a)$$

$$d\hat{Z}_s = \left(\frac{\theta - \hat{\theta}}{\eta\sqrt{1 - \rho^2}} - \frac{\mu(\theta) - \hat{\mu}}{\sigma} \frac{\rho}{\sqrt{1 - \rho^2}} \right) dt + dZ_s. \quad (1.11b)$$

Formally, $d\hat{Z}_Y$ and $d\hat{Z}_s$ are standard Brownian motions adapted to the firm's filtration. In less technical terms, they represent observable shocks that are part of the firm's information set. To see this, rewrite the stochastic processes for Y and s by substituting (1.11) into (1.3) and (1.6),

$$\frac{dY}{Y} = \hat{\mu}dt + \sigma d\hat{Z}_Y \quad (1.12a)$$

$$ds = \hat{\theta}dt + \eta\rho d\hat{Z}_Y + \eta\sqrt{1 - \rho^2}d\hat{Z}_s, \quad (1.12b)$$

so that Y and s are written using only observable information.

Finally, after some straightforward but tedious algebra, substitution of (1.11) into (1.9) yields

$$\begin{aligned} d\pi &= \kappa(\bar{\pi} - \pi)dt + \pi(1 - \pi) \left(\frac{\mu_h - \mu_l}{\sigma} \right) d\hat{Z}_Y \\ &\quad + \pi(1 - \pi) \left(\frac{\theta_h - \theta_l}{\eta\sqrt{1 - \rho^2}} - \frac{\mu_h - \mu_l}{\sigma} \frac{\rho}{\sqrt{1 - \rho^2}} \right) d\hat{Z}_s \\ &= \kappa(\bar{\pi} - \pi)dt + \pi(1 - \pi) \left(\omega_Y d\hat{Z}_Y + \omega_s d\hat{Z}_s \right), \end{aligned} \quad (1.13)$$

where, for ease of notation, I define the constants

$$\begin{aligned}\omega_Y &= \frac{\mu_h - \mu_l}{\sigma} \\ \omega_s &= \frac{\theta_h - \theta_l}{\eta\sqrt{1 - \rho^2}} - \frac{\mu_h - \mu_l}{\sigma} \frac{\rho}{\sqrt{1 - \rho^2}}.\end{aligned}$$

To summarize, while equations (1.3), (1.4), and (1.6) define an incomplete-information system, the filtering equations allow us to rewrite the stochastic processes in terms of observable parameters, transforming the equations into a complete information system.

Properties of the Learning Process

Several features of the learning process are worth pointing out. The first is that beliefs mean revert to $\bar{\pi}$ at rate κ . Indeed, given a prior $\pi(0)$, the conditional expectation of the posterior probability is given by

$$\begin{aligned}E[\pi(t)|\pi(0)] &= \pi(0)e^{-\kappa t} + \bar{\pi}(1 - e^{-\kappa t}) \\ &= \pi(0)e^{-(\lambda_{lh} + \lambda_{hl})t} + \left(\frac{\lambda_{lh}}{\lambda_{lh} + \lambda_{hl}}\right) \left(1 - e^{-(\lambda_{lh} + \lambda_{hl})t}\right)\end{aligned}\tag{1.15}$$

with

$$\lim_{t \rightarrow \infty} E[\pi(t)|\pi(0)] = \bar{\pi},$$

a weighted average of the prior and the long-run mean. This is the familiar form for mean-reverting stochastic processes. I derive this result in Appendix A.

The second property is that one statistic, π , characterizes both means and variances, which are given by

$$\begin{aligned}\hat{\theta} &= E[\theta|\pi] = \pi\theta_h + (1 - \pi)\theta_l \\ \hat{\mu} &= E[\mu(\theta)|\pi] = \pi\mu_h + (1 - \pi)\mu_l\end{aligned}$$

and

$$\begin{aligned}\text{Var}(\theta) &= E[(\theta - \hat{\theta})^2|\pi] = \pi(1 - \pi)(\theta_h - \theta_l)^2 \\ \text{Var}(\mu(\theta)) &= E[(\mu(\theta) - \hat{\mu})^2|\pi] = \pi(1 - \pi)(\mu_h - \mu_l)^2.\end{aligned}$$

Intuition for this result comes from the diffusion terms in equation (1.13). When π is close to zero or one, the diffusion terms, which are proportional to $\pi(1 - \pi)$, are small. In this case, beliefs, and therefore estimates of θ and $\mu(\theta)$, are precise, and any revision in beliefs in response to new information is small unless there is a large surprise. However, when beliefs are close to $\pi = \frac{1}{2}$, there is a large amount

of uncertainty about the state variable, and estimates are relatively imprecise. Even relatively little news about the state of the economy can result in a large revision of beliefs.

Finally, inspection of the diffusion terms in equation (1.13) suggest a dual role for demand volatility. In the standard theory of real options, higher volatility unambiguously leads to higher exercise thresholds and option values due to convexity of the payoff function. However, in the case of incomplete information, higher volatility has the effect of making it more difficult to learn about the economic environment. Mathematically, this is because volatility enters into the denominator of the diffusion terms; all else equal, higher volatility leads to smaller revisions in beliefs and slower resolution of uncertainty. Intuitively, this is because higher volatility leads to noisier observations which may mask the true state if, for example, noise cancels out any sustained periods of growth or decay in demand.

A reasonable guess then is that, for some regions of the parameter space, higher volatility may be associated with lower exercise thresholds, which I call a learning affect. All else equal, if delaying investment to observe demand conveys little to no information, the benefits to delaying may be small. I explore this possibility in Section 1.3.⁴

Steady-State Density

To explore the effects of volatility on learning, it is useful to examine the long-run stationary distribution of π . Denote $f(\pi, t)$ as the probability density function of π at any time t . Using standard results, the density function must satisfy the following partial differential equation,

$$\begin{aligned} \frac{\partial f}{\partial t} = & \frac{1}{2} \frac{\partial}{\partial \pi^2} \left[\pi^2 (1 - \pi)^2 \left(\frac{\mu_h - \mu_l}{\sigma} \right)^2 f(\pi, t) \right] \\ & - \frac{\partial}{\partial \pi} \left[(\lambda_{lh} + \lambda_{hl}) \left(\frac{\lambda_{lh}}{\lambda_{lh} + \lambda_{hl}} - \pi \right) f(\pi, t) \right]. \end{aligned}$$

This is known as the Kolmogorov forward equation. The steady-state, or stationary, distribution must also satisfy this equation, except that the derivative with respect to time is equal to zero. David (1997) and David and Veronesi (2002) show that the

⁴The same should be true for the noisy signals and the signal volatility η , except in this case the result should be unambiguous. Because noisy signals do not affect intrinsic value, there should only be a learning affect associated with η .

solution in the stationary case is

$$f(\pi) \propto \exp \left\{ \frac{-2(\lambda_{hl}\pi + \lambda_{lh}(1-\pi))}{\pi(1-\pi)(\omega_Y^2 + \omega_s^2)} \right\} \left(\frac{1-\pi}{\pi} \right)^{\frac{2(\lambda_{hl}-\lambda_{lh})}{\omega_Y^2 + \omega_s^2}} \quad (1.18)$$

$$\times \frac{1}{\pi^2(1-\pi)^2(\omega_Y^2 + \omega_s^2)}$$

Figure 1.1 plots the stationary distribution under four scenarios which highlight the role volatility and the transition parameters play in the resolution of uncertainty. Consider first the top two figures, which vary demand volatility while maintaining a symmetric transition density ($\lambda_{hl} = \lambda_{lh} = 0.25$). In the case with high volatility (top left), demand is a very imprecise indicator of the economic environment, so the revision in beliefs due to news is small. Noisy demand makes it difficult to filter out the state, so beliefs tend to be clustered around the mean. Hence, uncertainty resolves slowly, if at all.

Contrast this to the case with low volatility (top right). In this scenario, demand is a high-quality signal about the state of the economy. New signals convey a lot of information about the economy, so periods of high economic uncertainty are relatively rare. This translates into a bi-modal steady-state distribution - beliefs tend to be clustered near the two boundaries, and uncertainty resolves quickly.

The bottom two figures focus on the case of asymmetric parameters in the transition density, while holding volatility constant. In the bottom left figure, both λ_{hl} and λ_{lh} are large, with $\lambda_{hl} > \lambda_{lh}$. In this scenario, changes in the economic environment happen often, and recessions are more frequent and longer-lasting than expansions, which implies that the steady-state distribution will be skewed. This scenario is again one of slow learning. Recall that the rate of mean-reversion in beliefs is given by $\lambda_{hl} + \lambda_{lh}$. In this scenario, beliefs are reverting towards the mean at a high rate - the intuition is that it is difficult to filter out the state of the economy if it is changing frequently. Therefore, beliefs again tend to be clustered around the long-run mean, and uncertainty resolves slowly.

Finally, the last figure considers the case of asymmetric parameters with infrequent changes in the economic environment and $\lambda_{lh} > \lambda_{hl}$. This results in a skewed bi-modal distribution. The distribution is skewed because, in this scenario, expansions are more common than recessions, so beliefs tend to be more optimistic than pessimistic. It is bi-model because the relatively rare changes in the economic environment allow firms to form relatively precise beliefs - it is easier to learn when the economy is not always changing. Thus, in this scenario, uncertainty once again resolves quickly.

1.2.3 Project Valuation

Valuing the Cash Flows

Once the firm exercises its option to invest, the project generates per-period cash flows of $P(t)F(K) = Y(t)F(K)^{1-1/\gamma}$ in perpetuity. Firms are risk-neutral and discount cash flows at a constant rate r , with $r > \mu_h > \mu_l$:

$$V(Y, \pi; K, t) = E_t \left[\int_t^\infty e^{-r(u-t)} Y(u) F(K)^{1-\frac{1}{\gamma}} du \right] \quad (1.19)$$

I solve the expectation in (1.19) by recasting the present value function as the solution to a partial differential equation. Using a dynamic programming argument (see Section A for a derivation), the following valuation equation must hold for the present value of the cash flows:

$$rV dt = YF(K)^{1-\frac{1}{\gamma}} dt + E[dV] \quad (1.20)$$

In words, (1.20) says that the total return to owning the perpetual stream of cash flows (dividends $Y(t)F(K)^{1-1/\gamma}$ and expected capital gain $E[dV]/dt$) must equal the risk-free return rV . Using Ito's Lemma, $V(Y, \pi)$ obeys the stochastic process⁵

$$\begin{aligned} dV = & \left(\hat{\mu} Y V_Y + \frac{1}{2} \sigma^2 Y^2 V_{YY} + \kappa(\bar{\pi} - \pi) V_\pi \right. \\ & \left. + \frac{1}{2} \pi^2 (1 - \pi)^2 (\omega_Y + \omega_s)^2 V_{\pi\pi} + \pi(1 - \pi)(\mu_h - \mu_l) Y V_{Y\pi} \right) dt \\ & + \left(\sigma Y V_Y + \pi(1 - \pi) \omega_Y V_\pi \right) d\hat{Z}_Y + \pi(1 - \pi) \omega_s V_\pi d\hat{Z}_s \end{aligned} \quad (1.21)$$

Substituting the drift of (1.21) into (1.20) leads to the partial differential equation characterizing the present value of the cash flows,

$$\begin{aligned} rV = & YF(K)^{1-\frac{1}{\gamma}} + [\pi\mu_h + (1 - \pi)\mu_l] Y V_Y + \frac{1}{2} \sigma^2 Y^2 V_{YY} + \kappa(\bar{\pi} - \pi) V_\pi \\ & + \frac{1}{2} \pi^2 (1 - \pi)^2 (\omega_Y^2 + \omega_s^2) V_{\pi\pi} + \pi(1 - \pi) (\mu_h - \mu_l) Y V_{Y\pi}, \end{aligned} \quad (1.22)$$

with boundary condition $V(0, \pi; K, t) = 0$.

I solve the partial differential equation by the method of undetermined coefficients.

⁵The subscript notation denotes partial derivatives: $V_Y = \frac{\partial V}{\partial Y}$, $V_{YY} = \frac{\partial^2 V}{\partial Y^2}$, etc.

To begin, conjecture that the present value function takes the form

$$\begin{aligned} V(Y, \pi; K, t) &= H(\pi)Y(t)F(K)^{1-\frac{1}{\gamma}} \\ &= [A\pi + B(1 - \pi)]Y(t)F(K)^{1-\frac{1}{\gamma}} \end{aligned}$$

where A and B are coefficients to be determined. Substituting the conjecture into (1.22), simplifying, and collecting terms yields,

$$\begin{aligned} 0 = YF(K)^{1-\frac{1}{\gamma}} &\left([1 + A\kappa\bar{\pi} + B(\mu_l - r - \kappa\bar{\pi})] \right. \\ &\left. + \pi [A(\mu_h - r - \kappa) + B(r + \kappa - \mu_l)] \right) \end{aligned}$$

This equation must hold for all values of Y and π . In particular, for the equation to hold, the terms multiplying Y and π must equal zero. This implies that A and B must satisfy the system of equations

$$\begin{bmatrix} \kappa\bar{\pi} & \mu_l - r - \kappa\bar{\pi} \\ \mu_h - r - \kappa & r + \kappa - \mu_l \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

Therefore, the coefficients are

$$A = \frac{r - \mu_l + \lambda_{lh} + \lambda_{hl}}{(r - \mu_h)(r - \mu_l) + [\lambda_{lh}(r - \mu_h) + \lambda_{hl}(r - \mu_l)]} \quad (1.23a)$$

$$B = \frac{r - \mu_h + \lambda_{lh} + \lambda_{hl}}{(r - \mu_h)(r - \mu_l) + [\lambda_{lh}(r - \mu_h) + \lambda_{hl}(r - \mu_l)]} \quad (1.23b)$$

To understand the solution, it is useful to consider two limiting cases. In the first case, set $\lambda_{hl} = \lambda_{lh} = 0$. Then the solution for V simplifies to

$$V(Y, \pi; K, t) = \pi \frac{YF(K)^{1-\frac{1}{\gamma}}}{r - \mu_h} + (1 - \pi) \frac{YF(K)^{1-\frac{1}{\gamma}}}{r - \mu_l},$$

a probability-weighted average of two growing perpetuity formulas. Intuitively, when there are no changes in the economic environment, beliefs eventually converge to one or zero as firm learn the true value of θ . In this case, the value of future cash flows is simply the expected value over two possible outcomes.

At the other extreme, set $\lambda_{hl} = \lambda_{lh} = \lambda$. Then, as λ tends towards infinity, the solution is

$$\lim_{\lambda \rightarrow \infty} V(Y, \pi; K, t) = \frac{YF(K)^{1-\frac{1}{\gamma}}}{r - \frac{1}{2}(\mu_h + \mu_l)},$$

which does not depend on π . In this scenario, beliefs are constant at $\pi = 1/2$; the

state of the economy is changing every instant, and firms have no ability to identify the state.

Between these two extremes, the value of future cash flows is still linear in beliefs; λ_{hl} and λ_{lh} are parameters which control the sensitivity of project values to beliefs. When both parameters are small, expansionary and recessionary periods are highly persistent, so project values will be very sensitive to beliefs about the economic environment. The opposite case holds when both parameters are large, and the economy rapidly cycles between booms and busts.

Finally, substituting the solution for V into (1.21), the stochastic process for $V(Y, \pi)$ is given by

$$\frac{dV}{V} = \left[r - \frac{1}{H(\pi)} \right] dt + \left[\sigma + \frac{H_\pi}{H} \pi(1 - \pi)\omega_Y \right] d\hat{Z}_Y + \frac{H_\pi}{H} \pi(1 - \pi)\omega_s d\hat{Z}_s \quad (1.24)$$

so that $1/H(\pi)$ is the ‘dividend yield’ or opportunity cost of delaying investment. When firms forgo an investment project, they are passing up cash flows today to keep open the option to invest at a future date. Because H is linear and increasing in π , $1/H$ is monotonically decreasing in π . In other words, the opportunity cost of delaying investment decreases as beliefs become more optimistic. The intuition is that, in the high growth state, project valuations depend relatively more on future cash flows, while in the low growth state, project valuations depend relatively more on current cash flows. Therefore, delaying an investment is less costly when growth rates are high.

Furthermore, note that this framework endogenously generates stochastic volatility in project valuations, with volatility a quadratic function of beliefs. Because $H_\pi = A - B > 0$, volatility is increasing with economic uncertainty.

Valuing the Option to Invest

Standard real option models typically assume a fixed capital investment (i.e. K is fixed), but this assumption ignores that, in addition to choosing when to invest, firms have some flexibility in choosing how much to invest. In other words, K is a choice variable. Recognizing this flexibility, and denoting $G(Y, \pi)$ as the value of the firm’s investment option, the firm’s optimization problem is given by

$$G(Y, \pi) = \max_{K, \tau} E_0 \left[e^{-r\tau} (V(Y, \pi; K, \tau) - K) \right] \quad (1.25)$$

where the firm chooses both an optimal time to invest, τ , and an optimal level of investment, K .⁶

As with standard real option models, we can recast the optimal stopping problem

⁶This is just equation (1.1), where I have substituted in the solution for the present value of cash flows.

as an exercise threshold: Firms optimally delay investment until demand Y (or the output price P) reaches some threshold (which may vary with π). Once demand hits the threshold, firms invest the amount of capital that maximizes net present value, which follows from the first-order condition,

$$\frac{\partial V}{\partial K} = \left(1 - \frac{1}{\gamma}\right) [A\pi + B(1 - \pi)] Y^* F(K)^{-1/\gamma} F'(K) = 1. \quad (1.26)$$

Condition (1.26) simply requires that the marginal benefit of an additional unit of capital equals the marginal cost, and must hold if firms are investing to maximize NPV.

Due to the functional form of $F(K)$, the first-order condition is non-linear and must typically be solved numerically. However, in the case of a perfectly elastic demand curve ($\gamma \rightarrow \infty$), the first-order condition simplifies to

$$\frac{\partial V}{\partial K} = [A\pi + B(1 - \pi)] Y^* F'(K) = 1.$$

Using $F(K) = \log(1 + K)$ and imposing the additional constraint $K \geq 0$, the optimal capital choice conditional on reaching the investment threshold is

$$K^* = \max \{ [A\pi + B(1 - \pi)] Y^* - 1, 0 \}.$$

Owning an option to invest is analogous to owning a perpetual American call option on the investment project. At any point in time, a firm may either exercise their option and invest, or delay investment while retaining the option to invest at a future date. In the inaction region, where firms find it optimal to delay investment, the total return to owning the option must equal the risk-free return. (This is due to the assumption of risk-neutrality.) The option pays no dividends, i.e. the firm receives no cash flows from the project unless the option is exercised; total return is from capital gain only. Therefore, $rG dt = E[dG]$, the option analogue to the cash flow valuation equation in (1.20).

To obtain the expected capital gain, apply Ito's Lemma to $G(Y, \pi)$. Setting the drift term equal to $rG dt$ produces a partial differential equation for the value of the option to invest,

$$\begin{aligned} rG = & [\pi\mu_h + (1 - \pi)\mu_l] YG_Y + \frac{1}{2}\sigma^2 Y^2 G_{YY} + \kappa(\bar{\pi} - \pi)G_\pi \\ & + \frac{1}{2}\pi^2(1 - \pi)^2 (\omega_Y^2 + \omega_s^2) G_{\pi\pi} + \pi(1 - \pi) (\mu_h - \mu_l) YG_{Y\pi}. \end{aligned} \quad (1.27)$$

This is a free boundary problem (i.e. for each π we look for an optimal investment threshold Y^*), and the solution is subject to value-matching and smooth-pasting conditions.

The value-matching condition requires that, at the investment boundary, the value of the option equals the value from investing immediately. Therefore,

$$G(Y^*, \pi; \tau) = V(Y^*, \pi; K^*, \tau) - K^* = (A\pi + B(1 - \pi))Y^*F(K^*)^{1-\frac{1}{\gamma}} - K^*. \quad (1.28)$$

The smooth-pasting condition requires that, at the investment boundary, the value of the option and the intrinsic value meet tangentially. Therefore,

$$G_Y(Y^*, \pi; \tau) = V_Y(Y^*, \pi; K^*, \tau) \quad (1.29a)$$

$$G_\pi(Y^*, \pi; \tau) = V_\pi(Y^*, \pi; K^*, \tau). \quad (1.29b)$$

Finally, when the firm has the option to choose the scale of the project, investment must satisfy the marginal capital condition in (1.26).

The partial differential equation in (1.27) is similar to (1.22); in theory, the conditions in (1.26), (1.28), and (1.29) can be used to derive a solution. However, in this case, the free boundary problem rules out a closed-form solution, so numerical techniques are used.

1.3 Numerical Analysis

I solve the partial differential equation in (1.27) using a finite difference scheme; details of the grid and numerical procedure are in Appendix A. I solve the model under two scenarios. In the first, there are no signals, and the firm only observes demand. In the second, I allow for the firm to observe noisy signals. To improve computational efficiency, I also assume the demand curve is perfectly elastic, so that I can solve for the optimal capital allocation analytically.

1.3.1 Calibration to US GDP Growth

In order to study the model, I calibrate the model to US GDP growth. While I defer a discussion of the estimation method until Section 1.4, a brief discussion is warranted. Under the assumption that GDP is a reasonable proxy for aggregate demand, I estimate the parameters μ_h , μ_l , σ , λ_{hl} , and λ_{lh} , as well as a time series of filtered probabilities analogous to π , from the time series of GDP growth using the procedure for regime-switching models outlined in Chapter 22 of Hamilton (1994).

The calibration to GDP serves two purposes. First, it is always preferable to calibrate a model using reasonable parameters, so going to the data to find those parameters is a useful exercise. Second, the model will generate some predictions about the response of investment to beliefs about the economic environment; I use the filtered probabilities as regressors to test those predictions in Section 1.4.

Using quarterly data, I estimate annualized expansionary and recessionary growth rates of 4.1% and -1.8%, respectively, with annualized volatility of 3.1%. Additionally,

the estimate of λ_{hl} is 0.058, while the estimate of λ_{lh} is 0.302. These estimates imply that expansions last for about 17.3 quarters on average (or about 5.5 years), while recessions last for about 3.3 quarters on average (or just short of one year). Table 1.2 reports results of the estimation procedure, while the associated probabilities of being in the recessionary state are graphed in Figure 1.2. As a non-rigorous validation of the estimation and filtering procedure, the estimated probabilities line up well with NBER recession dates. Finally, I use $r = 5\%$ as the discount rate.

Benchmark Results

I first focus on the relationship between investment and growth rate uncertainty. Figure 1.4 plots the investment exercise threshold as a function of posterior beliefs for the case of both variable and fixed capital decisions. In terms of the partial differential equation (1.27), Figure 1.4 is the free boundary solution.

Consider first the top panel, which plots the exercise threshold when the investment scale of the project is a choice variable. The lower threshold is for the case when the firm only observes demand, while the upper threshold is for the case when the firm observes an additional noisy signal. Notice that the addition of a noisy signal increases the exercise boundary, leading to further delays in investment. This is because additional signals about demand improve the firm's ability to learn about the economy. Under the benchmark set of parameters drawn from U.S. GDP growth, the incentive to delay investing is a non-monotonic function of beliefs. Indeed, in the case with signals, the incentive to delay is highest when beliefs are near $\pi = 0.50$. Since the posterior probability π describes the probability that the economy is in a high-growth environment, uncertainty about beliefs is given by $\pi(1 - \pi)$, which is maximized at $\pi = 0.50$. Hence, if we believe that higher exercise boundaries translate into lower aggregate investment, Figure 1.4 suggests that investment is lowest almost precisely when uncertainty is highest, especially in the case of noisy signals.

Contrast this with the case where project scale is fixed. In this scenario, firms only choose when to invest, and the exercise threshold is plotted in the bottom panel of Figure 1.4. With fixed capital, the investment threshold declines monotonically as beliefs become more optimistic. The intuition for this result is that, while firms don't know the current economic environment, they do know that it can be only one of two cases, expansionary or recessionary. If demand is high enough such that firms would always invest if they knew the environment with certainty, then they would never delay; investing immediately is a dominant strategy.

What explains the markedly different results under the two cases? In the case with flexibility over project scale, there are two costs associated with exercising the investment option. One is the standard real option opportunity cost - by investing today, firms forgo the opportunity to exercise their option at a later date. When project scale is a choice variable, there is an additional cost: Firms risk irreversibly allocating a suboptimal level of capital to an investment project. As a result, the

benefit to delaying can be substantial during periods of high uncertainty if delaying allows the firm to make a more informed investment decision. This second cost is absent in the case of fixed capital decisions - investing immediately is a dominant strategy as long as demand is above the recessionary threshold as there is no risk of allocating the wrong amount of capital to a project.

To think about the fixed capital result another way, recall that the opportunity cost of delaying investment is $1/H(\pi)$, which is monotonic in beliefs since $H(\pi)$ is linear in beliefs. Since there is only an investment timing decision, and no investment scale decision, firms only face an invest-or-delay decision. Because the cost of delaying investment is monotonic in beliefs, the exercise threshold is also monotonic in beliefs.

The Effect of Volatility

Uncertainty about the economic environment resolves quickly when demand volatility is low. In this case, demand is highly informative about the true state of the economy, and learning is fast. Therefore, during periods of high uncertainty, there may be a large incentive to delay investment. Likewise, when volatility is high, demand is a very imprecise signal of the true state, and learning is slow. This suggests that the learning effects of high volatility may offset the usual convexity effects of volatility associated with higher expected payoffs.

Figure 1.5 plots the investment exercise threshold for various levels of volatility, and for both fixed and variable capital decisions. In the top panel, when capital is a choice variable, the effects of volatility appear to be striking. During periods of low economic uncertainty, when beliefs are close to either $\pi = 0$ or $\pi = 1$, the usual convexity effect dominates. Higher volatility appears to be associated with higher exercise thresholds - learning is relatively unimportant since there is little uncertainty about the economic environment.

However, as beliefs become less precise, the figure reveals a more complicated relationship between volatility and exercise thresholds. When demand volatility is low ($\sigma = 0.02$ or $\sigma = 0.03$ in the graph), beliefs about the economy can change quickly in response to new information, so delaying investment to learn can confer substantial benefits in the form of a more informed investment decision. For low enough levels of volatility, this effect can outweigh the convexity effect, so that lower volatility is associated with higher exercise thresholds. A similar result is found in Bernardo and Chowdhry (2002).

However, this effect appears to die out as volatility increases. Recall that higher volatility makes it more difficult to learn about the economy. Intuitively, there should be little benefit to delaying investment if information about the economy does not change. This is apparent for the higher volatility parameters ($\sigma = 0.04$ and $\sigma = 0.05$) in Figure 1.5, where the traditional relationship between volatility and exercise thresholds returns. As volatility increases, the benefits to learning diminish, so that the convexity effect outweighs the learning affect.

Finally, note that in the bottom figure, where project scale is fixed, the relationship between demand volatility and exercise thresholds appears to be monotonic. The numerical analysis suggests that the learning effects of volatility are closely linked to the additional cost associated with option exercise in the variable case. Delaying investment during periods of high economic uncertainty allows firms to make a more informed capital allocation decision. In a low-volatility environment, the benefits of delaying to learn can outweigh the opportunity costs associated with delaying exercise (the forgone cash flows). However, when capital is fixed and firms face only a timing decision, the cost of making a suboptimal capital investment are no longer present, and the economic benefits of delaying to learn are much smaller.⁷

The Effects of Signal Quality and Shift Intensity

Figure 1.6 plots the investment exercise threshold for various levels of signal quality η . The plots identify a negative relationship between noise and investment thresholds. As signal quality deteriorates (η increases), benefits to delaying investment diminish. Noisy signals are uninformative, so firms find them to be of little use in making their investment decision, while the delaying effect is strongest when signals are precise. The intuition closely resembles the learning effects of volatility - precise signals confer substantial learning benefits - except that there is no associated convexity effect since noise does not affect intrinsic value.

Figure 1.7 plots the investment exercise threshold for various levels of the transition density parameters, λ_{ht} and λ_{th} , and suggests that investment thresholds typically decline as shifts in the economic environment become more frequent. While qualitatively similar to the effect noise, the mechanism is different. Whereas volatility and noise affect the precision of beliefs, shift intensity affects average beliefs. When the economy is shifting frequently between expansions and recessions, beliefs are converging towards the mean at a higher rate; even if firms have precise beliefs about the economic environment today, they know that the environment is likely to change and become more uncertain. For high enough transition parameters, the effect of mean reversion dominates so that uncertainty effectively becomes long-lasting.

1.4 Empirical Analysis

With the numerical analysis of the investment boundary complete, let us now turn to the empirical analysis and a detailed explanation of the empirical procedure used to estimate the switching model parameters and associated probabilities.

⁷In the absence of closed-form solutions to the investment problem, a formal proof is unavailable. Therefore, any interpretation of the volatility effects in Figure 1.5 rely on the accuracy of the numerical procedure used to solve the problem.

1.4.1 Parameter Estimation

Estimation Procedure

I estimate the parameters of the switching model for U.S. GDP growth via maximum likelihood, using the method outlined in Chapter 22 of Hamilton (1994). Begin by applying Ito's Lemma to $y = \log Y$,

$$dy = \left[\pi\mu_h + (1 - \pi)\mu_l - \frac{1}{2}\sigma^2 \right] dt + \sigma d\hat{Z}_Y.$$

Then, conditional on θ , $y(t)$ is Normally distributed with density⁸

$$f_h = f(y(t)|\theta(t) = \theta_h) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y(t) - \mu_h + \frac{1}{2}\sigma^2}{\sigma} \right)^2 \right\}$$

$$f_l = f(y(t)|\theta(t) = \theta_l) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y(t) - \mu_l + \frac{1}{2}\sigma^2}{\sigma} \right)^2 \right\}.$$

Summing over the possible states θ_h and θ_l at $t - 1$, the density at time t is given by

$$f(y(t)) = (1 - \lambda_{hl})\pi_{t-1}f_h + \lambda_{hl}\pi_{t-1}f_l + \lambda_{lh}(1 - \pi_{t-1})f_h + (1 - \lambda_{lh})(1 - \pi_{t-1})f_l. \quad (1.31)$$

Finally, working from equation (1.31), the updated probabilities at time t are given by

$$\pi_t = \frac{(1 - \lambda_{hl})\pi_{t-1}f_h + \lambda_{lh}(1 - \pi_{t-1})f_h}{f(y(t))} \quad (1.32a)$$

$$1 - \pi_t = \frac{(1 - \lambda_{lh})(1 - \pi_{t-1})f_l + \lambda_{hl}\pi_{t-1}f_l}{f(y(t))}. \quad (1.32b)$$

Therefore, starting with an initial value π_0 , we can construct a series of conditional likelihoods and associated probabilities by iterating through equations (1.31) and (1.32).

Denote the parameter vector as Θ . Then the log-likelihood function is,

$$\mathcal{L}(\Theta) = \sum_{t=1}^T \log f(y(t))$$

which must be evaluated numerically.

⁸The reader should not confuse the posterior belief $\pi(t)$ with the mathematical constant π in the Normal distribution.

To calculate standard errors, I rely on the asymptotic distribution theory of maximum likelihood (see Campbell et al. (1997) for a reference), which states

$$\sqrt{T}(\hat{\Theta} - \Theta) \sim \mathcal{N}(0, \mathcal{I}^{-1}(\Theta))$$

where

$$\mathcal{I}(\Theta) = \lim_{T \rightarrow \infty} -E \left[\frac{1}{T} \frac{\partial^2 \mathcal{L}(\Theta)}{\partial \Theta \partial \Theta'} \right]$$

is the information matrix.

Rather than calculate second derivatives numerically, I rely on the information-matrix equality, which states that the information matrix can be estimated consistently from sample first derivatives as

$$\hat{\mathcal{I}} = \frac{1}{T} \sum_{t=1}^T \frac{\partial \log f(\hat{\Theta})}{\partial \Theta} \frac{\partial \log f(\hat{\Theta})'}{\partial \Theta} \quad (1.33)$$

The relevant derivatives are given in Appendix A.

Estimation Results

I fit the switching model to both quarterly and annual growth in real gross domestic product from 1951 to 2009. I source the data from the FRED (Federal Reserve Economic Data) database provided by the Federal Reserve Bank of St. Louis. To ease the comparison, I annualize quarterly growth rates. Table 1.2 presents the results.

Using quarterly data, I estimate annualized growth rates of 4.1% in the high state and -1.8% in the low state, with volatility of 3.1% per year. This compares to an unconditional mean and volatility of 3.2% and 3.8%, respectively, and the likelihood ratio test easily rejects the one-state model in favor of a two-state model. The estimates of λ_{hl} and λ_{lh} imply that expansions last for 17.3 quarters on average, while recessions last for 3.3 quarters on average (or about 5.5 years and just under 1 year).

The annual results give similar numbers for the length of regimes: Based on estimates for λ_{hl} and λ_{lh} , expansions last for about 5.2 years on average, while recessions last for about 1.4 years on average. However, the likelihood ratio test fails to reject the one-state model, and while the point estimate on μ_l is negative, it is statistically indistinguishable from zero. The quarterly estimate of λ_{lh} , which suggests recessions last for less than a year on average, provides a clue that annual data may be masking some short-lived recessions. Indeed, based on NBER recession dates, there were recessions in 1960, 1970, and 2001, each lasting for less than a year, that do not appear in annual GDP data. In other words, annual GDP growth rates were positive in each of these years. However, while these recessions do not appear in the annual data, they do appear in the estimated annual probabilities, as shown in the bottom half of

Figure 1.2.

Figure 1.3 plots the empirical and theoretical distribution of beliefs. The empirical density is simply the histogram of estimated probabilities, while the theoretical density is the predicted distribution according to equation (1.18). Empirically, the distribution of beliefs is bi-modal and skewed - beliefs are usually optimistic about the economy, with brief periods of pessimism, and periods of high uncertainty are rare. Although not visually obvious in the graphs, the estimated parameters imply a bi-modal theoretical distribution as well, with modes at $\pi = 0.10$ and $\pi = 0.98$ when fit to quarterly data, and modes at $\pi = 0.12$ and $\pi = 0.97$ when fit to annual data.

1.4.2 Empirical Analysis of Corporate Investment

Estimation Procedure

Given the theoretical predictions on the relationship between demand uncertainty and investment, as well as the evidence of high-growth and low-growth states of the US economy, a natural question to ask is the following: Do the filtered probabilities have explanatory power in explaining US investment behavior? Or are beliefs about the economy embedded within market values and Tobin's Q?

To explore these questions, I start with the following empirical specification:

$$\frac{I_{i,t}}{A_{i,t}} = \beta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{CF_{i,t}}{A_{i,t}} + \beta_3 Q_{i,t-1} + \beta_4 \pi_t + \beta_5 \pi_t^2 + \alpha_i + \epsilon_{i,t} \quad (1.34)$$

where $I_{i,t}$ is corporate investment, $A_{i,t}$ is the book value of assets, $CF_{i,t}$ is cash flow, $Q_{i,t}$ is a proxy for marginal Q, and π_t is the filtered belief that the economy is in an expansionary state. While this belief is common across all firms, all other regressors are firm-specific. The squared regressor captures any possible non-monotonic relationship between investment and beliefs. α_i is a firm-specific fixed effect, which captures unobserved heterogeneity across firms.

I also test a restricted form of (1.34), which combines beliefs and squared beliefs into one variable:

$$\frac{I_{i,t}}{A_{i,t}} = \beta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{CF_{i,t}}{A_{i,t}} + \beta_3 Q_{i,t-1} + \beta_4 \pi_t (1 - \pi_t) + \alpha_i + \epsilon_{i,t} \quad (1.35)$$

This restriction explicitly imposes $\beta_4 = -\beta_5$ in (1.34).

The regression equations in (1.34) and (1.35) present several empirical challenges for testing. The first is that the lagged dependent variable is cross-sectionally correlated with the fixed effect, as is the case in any dynamic panel regression. The idea is that large, unexplained shocks to investment go into the error term, which is the sum of the unobserved fixed effect and the random error $\epsilon_{i,t}$. All else equal, firms with large, positive investment shocks will have large fixed effects, while firms with

large, negative shocks to investment will have small fixed effects. This violates the assumptions required for consistency of the ordinary least squares estimator.

The standard method of removing fixed effects is to take a within-groups transformation of the data. In this transformation, group means are subtracted from the observations for each firm; since fixed effects are constant within groups, this removes the fixed effect. While this transformation is successful in eliminating the fixed effect, it has the unfortunate side effect of introducing correlation between the transformed lagged dependent variable and the transformed random error term. Unless there is a suitable instrument for lagged investment, within-groups estimation is inconsistent.⁹

An alternative method of removing fixed effects is the first differences estimator. In this transformation, the fixed effect is eliminated by taking first differences of all observations within groups. The resulting regression equation is

$$\Delta \frac{I_{i,t}}{A_{i,t}} = \beta_1 \Delta \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \Delta \frac{CF_{i,t}}{A_{i,t}} + \beta_3 \Delta Q_{i,t-1} + \beta_4 \Delta \pi_t + \beta_5 \Delta \pi_t^2 + \Delta \epsilon_{i,t}. \quad (1.36)$$

where Δ is the first differences operator.

This approach still introduces correlation between the transformed lagged dependent variable and the transformed error term, since both transformed variables contain $t - 1$ terms:

$$\Delta \frac{I_{i,t-1}}{A_{i,t-1}} = \frac{I_{i,t-1}}{A_{i,t-1}} - \frac{I_{i,t-2}}{A_{i,t-2}} \quad (1.37a)$$

$$\Delta \epsilon_{i,t} = \epsilon_{i,t} - \epsilon_{i,t-1}. \quad (1.37b)$$

The correlation arises because, from equation (1.34), investment at $t - 1$ is correlated with the error term at $t - 1$. However, as long as the transformed error term is serially uncorrelated, deeper lags of investment, which contain no $t - 1$ terms, will be uncorrelated with the error. Therefore, unlike in the within groups transformation, deeper lags are potentially valid instruments for lagged investment. This approach to using deeper lags as instruments in dynamic panels is due to Arrelano and Bond (1991).

An additional empirical challenge is that cash flow, used as a firm control in (1.34), is simultaneously determined with investment. It too will be correlated with the error term in (1.34) and (1.36) due to un-modeled operating decisions which jointly affect both investment and cash flow. Again, we require a suitable instrument for cash flow. Appealing to the logic used above for investment, lagged values of cash flow are potentially suitable, if not economically meaningful, instruments if the error term is i.i.d. Brown and Petersen (2009) also take this approach to instrumenting cash flow.

To summarize, I estimate the parameters in (1.34) using two-stage least squares on first differences. The excluded instrument set contains the second, third, and

⁹See Roodman (2006) and Bond (2002) for a discussion of this point.

fourth lags of cash flow, and the third and fourth lags of investment. With more instruments than endogenous regressors, the system is over-identified; this allows us to use the χ^2 test of over-identifying restrictions to test the orthogonality conditions. If the χ^2 statistic is sufficiently large, we reject the hypothesis that the instruments are uncorrelated with the error term.

The included instrument set contains Q , π , and π^2 . In the first stage regressions I fit cash flow and lagged investment to the entire set of included and excluded instruments, while in the second stage I regress investment on the included instruments and the fitted values of lagged investment and cash flow. The first stage F-test of excluded instruments tests the joint significance of the instrumental variables. Taken together, the F-test and χ^2 statistic provide two useful metrics in determining the validity of the instruments.¹⁰

Data Description

The empirical study utilizes annual data on US firms from 1964 to 2008. Firm data come from the COMPUSTAT database; I use 1964 as the beginning date due to limited data availability prior to 1964. In most cases I follow Lewellen and Lewellen (2010) when constructing the relevant variables.¹¹

Table 1.1 contains a list and description of the relevant variables from COMPUSTAT. The first step is filtering down the database to the set of relevant firms. First, following convention, I classify firm-years according to the fiscal year-end month. If a firm's fiscal year ends in January through May, that observation is classified to the prior year. For example, Circuit City's fiscal year ended in March prior to 1979, and February thereafter. Therefore, the COMPUSTAT record dated March 31, 1977 is assigned to the year 1976.

Next, I keep only firms headquartered within the United States, as well as only those with ISO or native currency equal to the U.S. dollar. Finally, I remove any firms with SIC codes ranging from 6000 to 6799 ('Fire, Insurance, and Real Estate' division) or 9100 to 9999 ('Public Administration' division).

I consider two measures of investment. The first is simply capital expenditures, from the statement of cash flows. Following Lewellen and Lewellen (2010), I also consider a measure of the investment which is the sum of capital expenditures and cash spent on acquisitions (again, from the statement of cash flows).

Cash flow is defined as operating income before depreciation. However, this variable is not always reported, so I rely on some accounting relationships to fill in missing values where possible. If operating income is missing, I then define cash flow as income before extraordinary items plus extraordinary items and depreciation and

¹⁰I implement the two-stage regression procedure in Stata using the `xtivreg2` command authored by Schaffer (2007).

¹¹Note that, because of the use of both first differences and up to four lags of investment and cash flow as instruments, the first observation in the estimation is 1969.

amortization, both from the statement of cash flows. Occasionally, depreciation and amortization is missing as well, so I replace this with depreciation and amortization from the income statement where possible.

I define net assets as book value of total assets less non-debt current liabilities, and calculate net debt as total liabilities less non-debt current liabilities. I fill in missing net asset values with the sum of net debt and book equity, where possible. Likewise, I fill in missing debt values with the sum of long-term debt, short-term debt, and other liabilities.

Market value of equity is the closing share price times the number of common shares outstanding, and I define Q as the sum of market equity and net debt, scaled by net assets. I then drop any observations with either a market value of equity equal to zero, or a book value of equity less than or equal to zero, and also any observations with missing values for net assets, investment, cash flow, or Q . This results in a final data set with 158,491 firm-year observations.

Finally, since both cash flow and investment are flow variables that occur throughout the year, I scale each by the average of beginning-year and ending-year net assets. Furthermore, to reduce the impact of outliers and associated measurement error, I Winsorize investment, cash flow, and Q at the 1st and 99th percentiles. This adjustment, common in the literature, converts all values in the upper and lower tails of the distribution to the 1st and 99th percentile values.

Estimation Results

I estimate (1.34) using annual data for all COMPUSTAT firms, and also split the sample into manufacturing and non-manufacturing firms. To review, the relevant regression equation is

$$\frac{I_{i,t}}{A_{i,t}} = \beta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{CF_{i,t}}{A_{i,t}} + \beta_3 Q_{i,t-1} + \beta_4 \pi_t + \beta_5 \pi_t^2 + \alpha_i + \epsilon_{i,t} \quad (1.38)$$

where π measures the estimated probability that the economy is in an expansionary environment. I consider two measures of beliefs. The first is simply the probability of economic expansion estimated from annual data. However, I also estimate the regression using average quarterly beliefs over the year, for two reasons. First, since beliefs evolve dynamically and investment takes place throughout the year, a beginning-year annual belief can become stale rather quickly. Second, the likelihood ratio tests reject the two-state model using annual data, but fail to reject using quarterly data. Nevertheless, the regressions using annual beliefs perform better in terms of testing the over-identifying restrictions, perhaps because smoothness induced by averaging quarterly data.

I also estimate a restricted form of the regression,

$$\frac{I_{i,t}}{A_{i,t}} = \delta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \delta_2 \frac{CF_{i,t}}{A_{i,t}} + \delta_3 Q_{i,t-1} + \delta_4 \pi_t (1 - \pi_t) + \alpha_i + \epsilon_{i,t} \quad (1.39)$$

which imposes the restriction $\beta_4 = -\beta_5$ in (1.39). I do this because $\pi(1 - \pi)$ is the measure of economic uncertainty in the theoretical framework, but also because of concerns about multi-collinearity. Because estimated beliefs are typically close to zero or one, squared beliefs are also typically close to zero or one. As we shall see, in eleven out of twelve cases, the point estimate of β_3 is negative and the point estimate of β_4 is positive. Although this result is consistent with the theoretical framework, the alternating coefficients could be indicative of two highly collinear variables.

Turning now to the regression results, Table 1.4 presents results using 1) capital expenditures as the measure of investment, and 2) estimated annual beliefs. The point estimates of β_4 and β_5 are -0.016 and 0.022, respectively, which implies a quadratic relationship between investment and beliefs. Starting out at $\pi = 0$, investment initially declines as beliefs become more optimistic - even though the probability of an expansionary environment is increasing, economic uncertainty is also increasing. However, after reaching some threshold, investment begins to increase as beliefs become more optimistic and uncertainty falls. The pattern is qualitatively the same for the manufacturing and non-manufacturing sub-samples, although the χ^2 test rejects the model for the manufacturing sector. All coefficients are significant at the 5% level, with all but one significant at the 1% level, and the F-tests of instrument significance reject the null of joint insignificance.

The restricted estimates of δ_4 are -0.030 for all firms, -0.027 for manufacturing firms, and -0.036 for non-manufacturing firms, again implying that investment rates systematically decrease as economic uncertainty increases. However, Wald tests easily reject the restrictions, suggesting that the unrestricted regressions provide a better fit to the data.

Although the patterns of the estimated coefficients are the same, regressions using the average of quarterly beliefs (reported in Table 1.5) are somewhat less successful overall in terms of χ^2 tests, which reject the over-identifying restrictions for all firms and the manufacturing sub-sample. The χ^2 test fails to reject the model for the non-manufacturing sub-sample, but the estimates of β_4 and β_5 are statistically indistinguishable from zero. However, the estimate of δ_4 in the restricted regression is significant, which suggests that multi-collinearity in beliefs and squared beliefs are combining with the smaller sample size to mask any relationship between investment and beliefs for non-manufacturing firms. One notable difference when using average quarterly beliefs is the Wald test, which fails to reject the hypothesis that $\beta_4 = -\beta_5$ in all three cases.

Tables 1.6 and 1.7 present regression results when the measure of investment is the sum of capital expenditures and cash spent on acquisitions, and the results are broadly

consistent with the previous results, except that the magnitudes are somewhat larger. For all firms using annual beliefs, the point estimates of β_4 and β_5 are -0.024 and 0.030, respectively, and the Wald test rejects the hypothesized restriction on the coefficients. Furthermore, the χ^2 test fails to reject the over-identifying restrictions in every case, though again the point estimates in the non-manufacturing are statistically zero when using average quarterly beliefs.¹²

Figure 1.8 plots the expected level of investment conditional on beliefs, in both the restricted and unrestricted cases. In the top panel, the measure of investment is capital expenditures, while the bottom panel includes acquisitions. In all cases, the figures display the quadratic relationship between investment and beliefs described above, which are consistent with the theoretical results when firms have flexibility over the scale of their investments. (Recall that higher investment thresholds translate into lower aggregate investment.)

In summary, the empirical findings suggest that uncertainty about the economic environment indeed plays a role in corporate investment decisions, consistent with the theoretical framework. The parameters of the switching model estimated from U.S. GDP growth imply a distribution of beliefs that is bi-modal, with quick resolution of uncertainty, and exercise thresholds that are approximately quadratic in beliefs (or linear in uncertainty). Using probabilities estimated from time series of GDP growth, the empirical relationship between corporate investment and beliefs is quadratic, which is consistent with the theoretical model with flexible capital allocation.

1.5 Concluding Remarks

Economic growth and demand uncertainty is important for corporate investment decisions if the demand environment affects the cash flows generated by an investment project and investment is irreversible. Guided by this intuition, I develop a theoretical framework which incorporates growth rate uncertainty into a firm's investment decision.

The model developed in this chapter suggests that firms will decrease investment as growth uncertainty increases provided that 1) firms have flexibility over the scale of their investments, and 2) uncertainty resolves quickly. Uncertainty will typically resolve quickly as long as demand volatility is relatively low, signals about demand are relatively precise, and changes in the economic environment happen infrequently.

Empirically, real gross domestic product displays precisely the properties necessary for uncertainty to resolve quickly. Although we do not observe noisy signals, volatility

¹²A caveat about the results: The estimated standard errors do not account for the fact the the probabilities used in the regressions are themselves estimates and subject to error. As a consequence, the reported standard errors are potentially too small, and the statistical significance is perhaps less remarkable than reported.

is sufficiently low and shifts in the economy are sufficiently infrequent to generate the bi-modal distribution of beliefs that is consistent with quick resolution of uncertainty. Therefore, we might expect corporate investment rates to respond systematically to beliefs about the economy.

Indeed, estimated probabilities from GDP growth are useful in explaining corporate investment. The estimated relationship between investment and beliefs is quadratic; correspondingly, the estimated relationship between investment and uncertainty is linear. In sum, uncertainty about the state of the economy appears to be important for corporate investment, both theoretically and empirically.

1.6 Figures and Tables

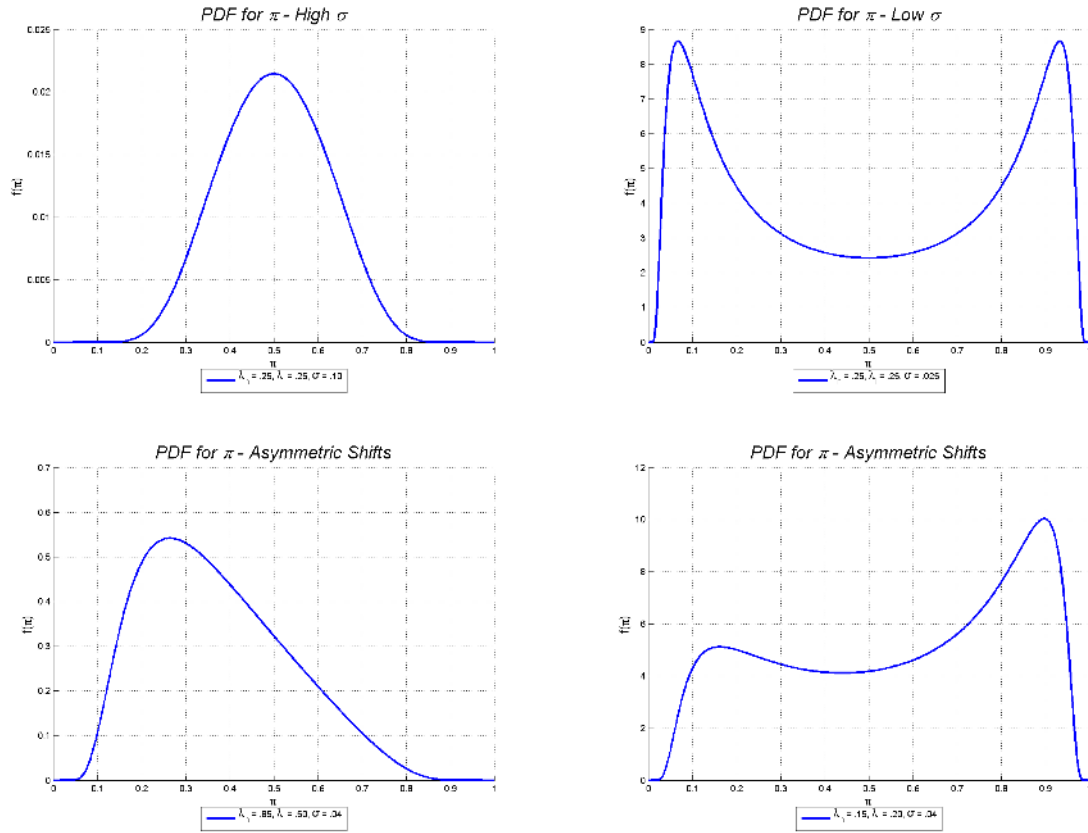


Figure 1.1: Steady-state probability density for π given by equation (1.18) when learning is slow (left), and when learning is fast (right), under the assumption that no additional signals are observed. In the top panel, probabilities of regime shifts are symmetric, while in the bottom panel they are asymmetric. Demand growth is 4% in the high state and -1% in the low state.

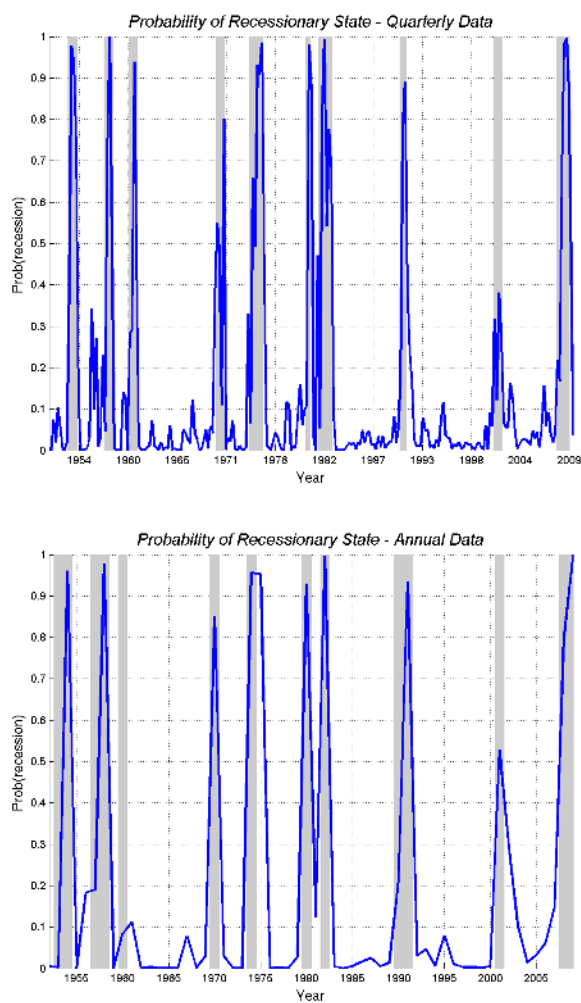


Figure 1.2: Probability that the economy is in a recessionary state based on US GDP growth, 1951 - 2009. The shaded bars represent recessions as defined by NBER. I estimate probabilities using the procedure outlined in Chapter 22 of Hamilton (1994). Table 1.2 provides the associated parameter estimates of the regime-switching model.

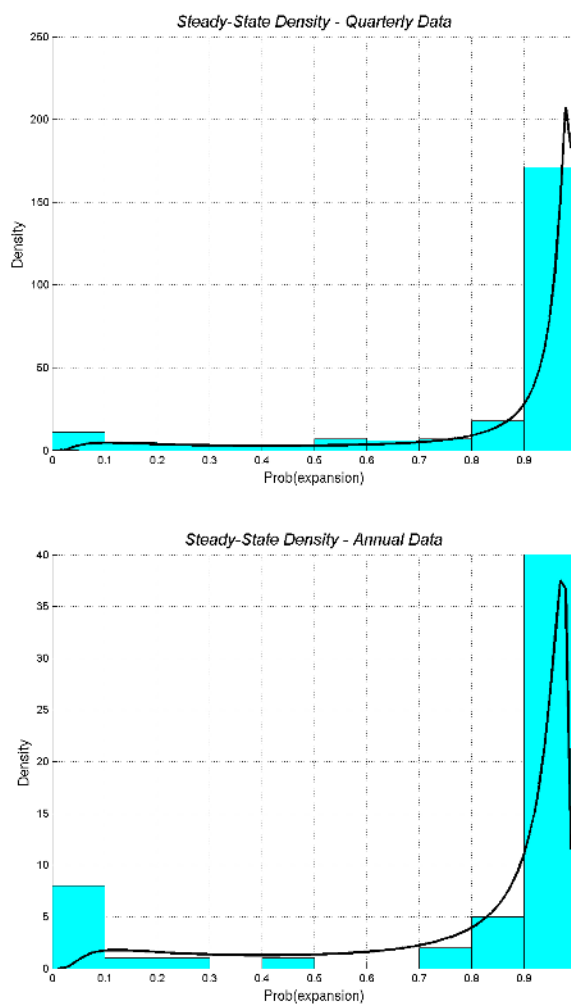


Figure 1.3: Empirical and theoretical densities of the probability that the economy is in a expansionary state based on US GDP growth, 1951 - 2009. I estimate probabilities using the procedure outlined in Chapter 22 of Hamilton (1994). The theoretical density, calculated with estimated parameters, is given in equation (1.18).

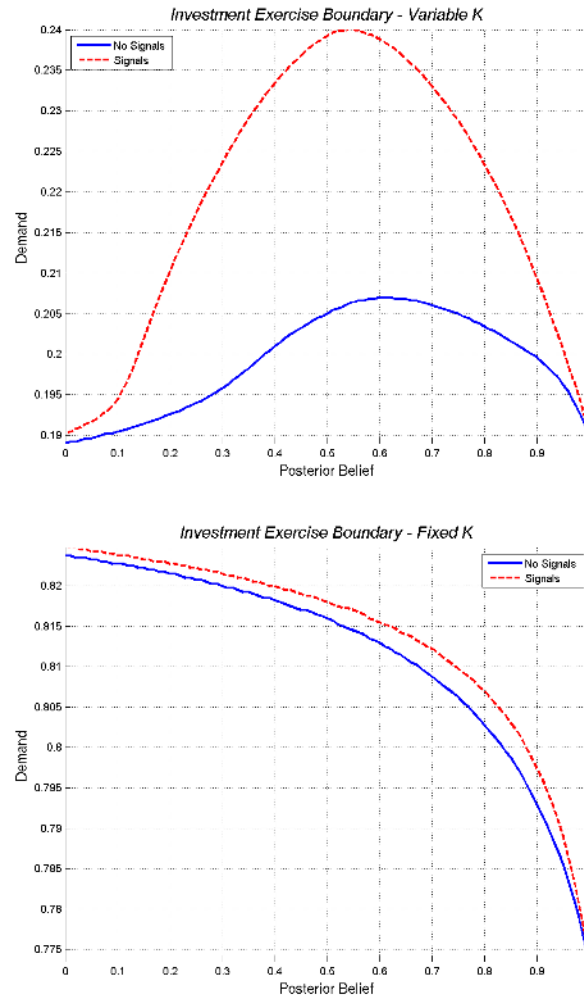


Figure 1.4: Free boundary solution to the PDE in equation (1.27) when investment scale is a choice variable (top), and when investment scale is fixed (bottom). The PDE solution is based on quarterly parameter estimates of the regime-switching model given in Panel A of Table 1.2. In the case with noisy signals, $\eta = 0.40$.

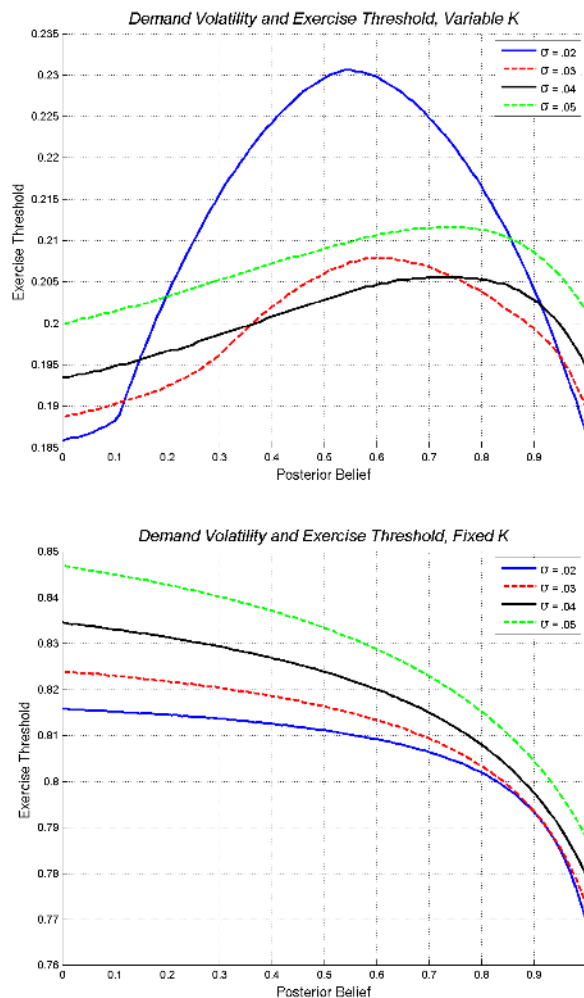


Figure 1.5: Free boundary solution to the PDE in equation (1.27) for varying levels of demand volatility σ when investment scale is a choice variable (top), and when investment scale is fixed (bottom). The PDE solution is based on quarterly parameter estimates of the regime-switching model given in Panel A of Table 1.2.

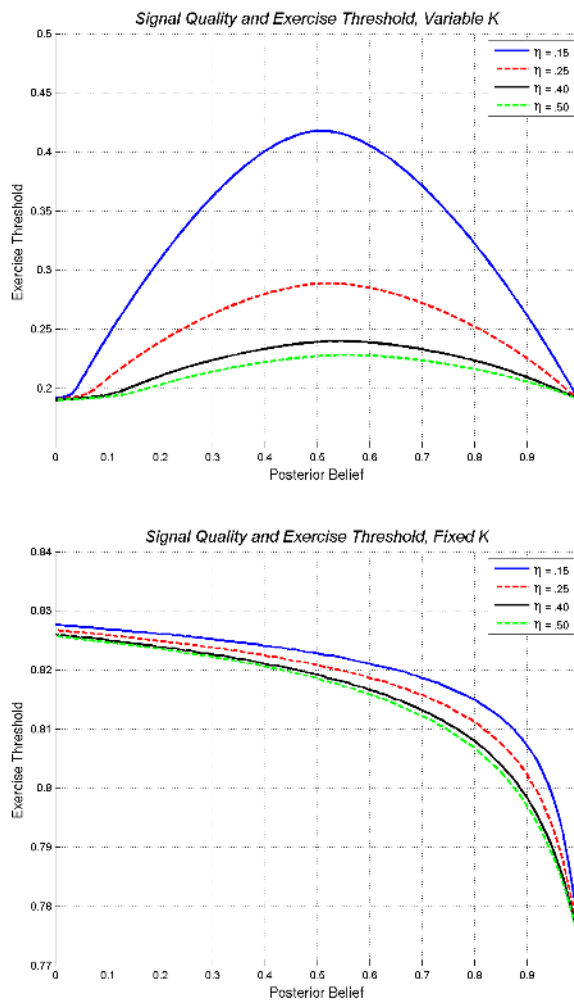


Figure 1.6: Free boundary solution to the PDE in equation (1.27) for varying levels of noise quality η when investment scale is a choice variable (left), and when investment scale is fixed (right). The PDE solution is based on quarterly parameter estimates of the regime-switching model given in Panel A of Table 1.2.

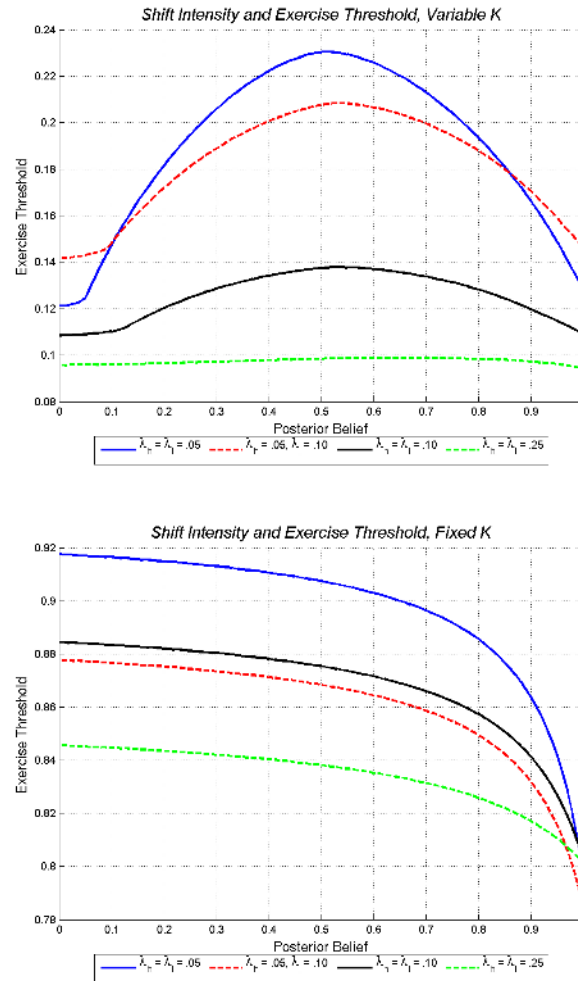


Figure 1.7: Free boundary solution to the PDE in equation (1.27) for varying economic transition parameters when investment scale is a choice variable (left), and when investment scale is fixed (right). The PDE solution is based on quarterly parameter estimates of the regime-switching model given in Panel A of Table 1.2.

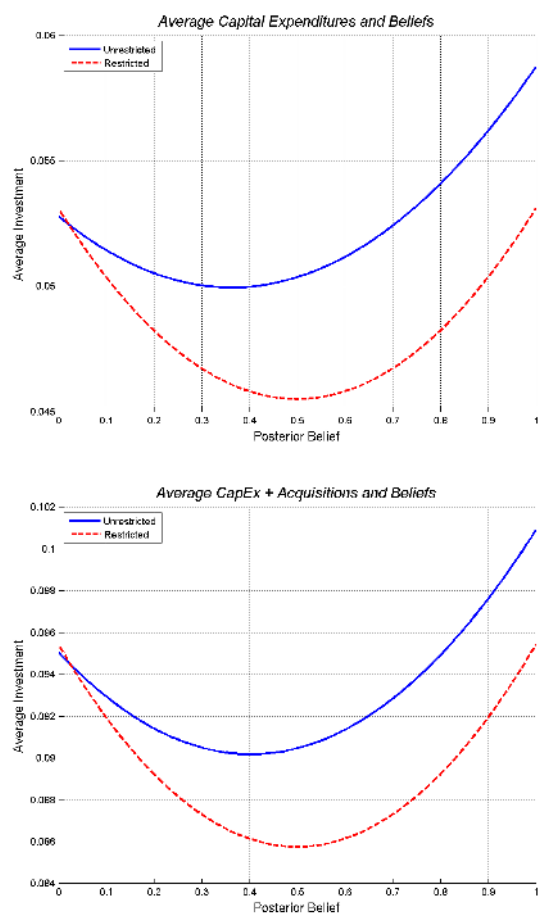


Figure 1.8: This figure displays the expected level of investment conditional on beliefs for both restricted and unrestricted cases, where investment is measured as capital expenditures (top) and the sum of capital expenditures and acquisitions (bottom). Capital expenditure plots are based on regression estimates in Table 1.4, while capital expenditure plus acquisition plots are based on regression estimates in Table 1.6.

Table 1.1: Description of COMPUSTAT variables

Data Code	Descriptor
gvkey	Global Company Key
loc	Current ISO Country Code - Headquarters
sic	Standard Industry Classification Code
datadate	Data Date
fy	Fiscal Year-end Month
curcd	ISO Currency Code
curncd	Native Currency Code
aqc	Acquisitions
at	Assets - Total
capx	Capital Expenditures
cshe	Common Shares Outstanding
dlc	Debt in Current Liabilities - Total
dltt	Long-Term Debt - Total
dp	Depreciation and Amortization
dpc	Depreciation and Amortization (Cash Flow)
ibc	Income Before Extraordinary Items (Cash Flow)
lct	Current Liabilities - Total
lo	Liabilities - Other - Total
lt	Liabilities - Total
oiadp	Operating Income After Depreciation
oibdp	Operating Income Before Depreciation
prcc_c	Price Close - Annual - Calendar
seq	Shareholders' Equity - Total
xidoc	Extraordinary Items and Discontinued Operations (Cash Flow)

Table 1.2: **Regime-Switching Parameters for GDP Growth**

This table presents maximum likelihood estimates of the two-state model for GDP growth from 1951 to 2009, for both quarterly and annual data. For comparison purposes, quarterly growth rates are annualized. $\mathcal{L}(\Theta)$ is the maximized value of the log-likelihood function, and LR is the likelihood ratio statistic comparing the two-state model to a one-state model. Maximum likelihood standard errors are in parentheses, and p-values for the likelihood ratio test come from Table 1A in Garcia (1998).

Panel A: Two-State Model							
	μ_h	μ_l	σ	λ_{hl}	λ_{lh}	$\mathcal{L}(\Theta)$	T
Quarterly	0.041 (0.004)	-0.018 (0.006)	0.031 (0.002)	0.058 (0.102)	0.302 (0.153)	450.340	235
Annual	0.039 (0.003)	-0.001 (0.007)	0.015 (0.002)	0.192 (0.147)	0.701 (0.099)	141.753	58
Panel B: One-State Model							
	μ	σ	$\mathcal{L}(\Theta)$	LR	p-value	T	
Quarterly	0.032 (0.003)	0.038 (0.001)	436.818	27.045	0.000	235	
Annual	0.031 (0.003)	0.022 (0.002)	138.986	5.535	0.300	58	

Table 1.3: Summary Statistics - Investment, Cash Flow, and Tobin's Q
 This table presents summary statistics for corporate financial data from the COMPUSTAT database, before and after adjusting for outliers. Capital expenditures, acquisitions, and cash flow are scaled by the average of beginning and ending book value of assets, and Tobin's Q is calculated as the market value of equity plus debt, scaled by book assets. To adjust for outliers, variables are Winsorized at the 1st and 99th percentiles.

Panel A: Pre-Winsor				
	Capital Expenditures	CapEx + Acquisitions	Cash Flow	Tobin's Q
Mean	0.097	0.120	0.107	2.442
Median	0.066	0.079	0.156	1.327
Standard Deviation	0.118	0.147	0.414	20.351
Skewness	18.347	10.827	-39.375	230.409
Kurtosis	1,919.142	801.461	4,002.407	68,717.710
Minimum	-0.209	-1.526	-55.628	0.003
Maximum	15.569	15.569	3.670	6,530.038
1st percentile	0.000	0.000	-1.102	0.421
99th percentile	0.537	0.704	0.603	15.547
Panel B: Post-Winsor				
	Capital Expenditures	CapEx + Acquisitions	Cash Flow	Tobin's Q
Mean	0.094	0.118	0.115	2.102
Median	0.066	0.079	0.156	1.327
Standard Deviation	0.095	0.125	0.253	2.309
Skewness	2.252	2.329	-2.205	3.552
Kurtosis	9.161	9.485	10.345	17.956
Minimum	0.000	0.000	-1.102	0.421
Maximum	0.537	0.704	0.603	15.547
1st percentile	0.000	0.000	-1.102	0.421
99th percentile	0.537	0.704	0.603	15.547

Table 1.4: **IV Regressions - Capital Expenditures and Annual Beliefs**

This table presents second stage IV regressions of corporate investment on cash flow, lagged investment, Tobin's Q, and estimated beliefs about the economy. The second-stage regression equation is:

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{CF_{i,t}}{A_{i,t}} + \beta_3 Q_{i,t-1} + \beta_4 \pi_t + \beta_5 \pi_t^2 + \epsilon_{i,t}$$

where the dependent variable is capital expenditures scaled by assets, and π is the belief that the economy is in an expansionary state. The Wald test is a test of the restriction that $\beta_4 + \beta_5 = 0$; the restricted regression imposes this restriction. Cash flow and lagged investment are endogenous controls and dependent variables in unreported first-stage regressions. The set of first stage instruments includes cash flow lagged 2, 3, and 4 years, and investment lagged 3 and 4 years; F-statistics test the joint significance of the instruments. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. All regressions include firm fixed effects; robust standard errors, clustered by firm, are shown in parentheses. The χ^2 statistic with three degrees of freedom (J Test) tests the over-identifying restrictions.

	All Firms		Manufacturing		Non-Manufacturing	
	Unrestricted	Restricted	Unrestricted	Restricted	Unrestricted	Restricted
Cash Flow _t	0.146*** (0.019)	0.150*** (0.019)	0.126*** (0.019)	0.132*** (0.019)	0.166*** (0.037)	0.168*** (0.037)
Investment _{t-1}	0.297*** (0.037)	0.296*** (0.037)	0.222*** (0.037)	0.221*** (0.037)	0.378*** (0.069)	0.377*** (0.069)
Q	0.004*** (0.000)	0.004*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.005*** (0.001)	0.005*** (0.001)
π	-0.016*** (0.004)		-0.011** (0.005)		-0.023*** (0.007)	
π^2	0.022*** (0.004)		0.017*** (0.005)		0.029*** (0.007)	
$\pi \times (1 - \pi)$		-0.030*** (0.004)		-0.027*** (0.004)		-0.036*** (0.007)
CF Stage 1 F-stat	141.25	114.23	79.56	79.66	41.24	41.22
p-value	0.000	0.000	0.000	0.000	0.000	0.000
I Stage 1 F-stat	233.60	232.520	189.35	189.18	81.83	81.40
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test ($\chi^2_{(1)}$)	81.10		72.36		20.41	
p-value	0.000		0.000		0.000	
J Test ($\chi^2_{(3)}$)	5.836	5.723	8.085	7.655	1.658	1.689
p-value	0.120	0.126	0.044	0.054	0.646	0.639
# Obs.	92,117	92,117	52,333	52,333	39,784	39,784
# Clusters	8,580	8,580	4,448	4,448	4,132	4,132

Table 1.5: **IV Regressions - Capital Expenditures and Average Quarterly Beliefs**

This table presents second stage IV regressions of corporate investment on cash flow, lagged investment, Tobin's Q, and estimated beliefs about the economy. The second-stage regression equation is:

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{CF_{i,t}}{A_{i,t}} + \beta_3 Q_{i,t-1} + \beta_4 \pi_t + \beta_5 \pi_t^2 + \epsilon_{i,t}$$

where the dependent variable is capital expenditures scaled by assets, and π is the belief that the economy is in an expansionary state. The Wald test is a test of the restriction that $\beta_4 + \beta_5 = 0$; the restricted regression imposes this restriction. Cash flow and lagged investment are endogenous controls and dependent variables in unreported first-stage regressions. The set of first stage instruments includes cash flow lagged 2, 3, and 4 years, and investment lagged 3 and 4 years; F-statistics test the joint significance of the instruments. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. All regressions include firm fixed effects; robust standard errors, clustered by firm, are shown in parentheses. The χ^2 statistic with three degrees of freedom (J Test) tests the over-identifying restrictions.

	All Firms		Manufacturing		Non-Manufacturing	
	Unrestricted	Restricted	Unrestricted	Restricted	Unrestricted	Restricted
Cash Flow _t	0.153*** (0.019)	0.153*** (0.019)	0.137*** (0.019)	0.137*** (0.019)	0.168*** (0.037)	0.168*** (0.037)
Investment _{t-1}	0.301*** (0.038)	0.301*** (0.038)	0.224*** (0.037)	0.224*** (0.037)	0.384*** (0.070)	0.384*** (0.070)
Q	0.004*** (0.000)	0.004*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.005*** (0.001)	0.005*** (0.001)
π	-0.017** (0.008)		-0.022** (0.009)		-0.009 (0.014)	
π^2	0.016*** (0.006)		0.018*** (0.006)		0.012 (0.010)	
$\pi \times (1 - \pi)$		-0.014*** (0.003)		-0.010*** (0.004)		-0.020*** (0.006)
CF Stage 1 F-stat	114.50	114.39	80.89	80.85	41.01	40.98
p-value	0.000	0.000	0.000	0.000	0.000	0.000
I Stage 1 F-stat	228.85	228.22	185.00	184.77	80.48	80.14
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test ($\chi^2_{(1)}$)	0.19		1.98		0.67	
p-value	0.667		0.160		0.414	
J Test ($\chi^2_{(3)}$)	6.339	6.309	8.333	8.113	1.895	1.893
p-value	0.096	0.098	0.040	0.044	0.595	0.595
# Obs.	92,117	92,117	52,333	52,333	39,784	39,784
# Clusters	8,580	8,580	4,448	4,448	4,132	4,132

Table 1.6: IV Regressions - Capital Expenditures and Acquisitions, Annual Beliefs

This table presents second stage IV regressions of corporate investment on cash flow, lagged investment, Tobin's Q, and estimated beliefs about the economy. The second-stage regression equation is:

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{CF_{i,t}}{A_{i,t}} + \beta_3 Q_{i,t-1} + \beta_4 \pi_t + \beta_5 \pi_t^2 + \epsilon_{i,t}$$

where the dependent variable is capital expenditures scaled by assets, and π is the belief that the economy is in an expansionary state. The Wald test is a test of the restriction that $\beta_4 + \beta_5 = 0$; the restricted regression imposes this restriction. Cash flow and lagged investment are endogenous controls and dependent variables in unreported first-stage regressions. The set of first stage instruments includes cash flow lagged 2, 3, and 4 years, and investment lagged 3 and 4 years; F-statistics test the joint significance of the instruments. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. All regressions include firm fixed effects; robust standard errors, clustered by firm, are shown in parentheses. The χ^2 statistic with three degrees of freedom (J Test) tests the over-identifying restrictions.

	All Firms		Manufacturing		Non-Manufacturing	
	Unrestricted	Restricted	Unrestricted	Restricted	Unrestricted	Restricted
Cash Flow _t	0.268*** (0.041)	0.273*** (0.041)	0.195*** (0.042)	0.203*** (0.042)	0.324*** (0.080)	0.325*** (0.079)
Investment _{t-1}	0.401*** (0.074)	0.401*** (0.074)	0.228*** (0.079)	0.225*** (0.079)	0.560*** (0.125)	0.560*** (0.125)
Q	0.008*** (0.001)	0.008*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.012*** (0.001)	0.012*** (0.001)
π	-0.024*** (0.008)		-0.020* (0.010)		-0.039*** (0.014)	
π^2	0.030*** (0.008)		0.028*** (0.010)		0.042*** (0.013)	
$\pi \times (1 - \pi)$		-0.039*** (0.008)		-0.041*** (0.010)		-0.045*** (0.013)
CF Stage 1 F-stat	113.97	113.94	79.72	79.82	41.29	41.26
p-value	0.000	0.000	0.000	0.000	0.000	0.000
I Stage 1 F-stat	109.03	108.30	68.42	68.16	46.41	46.18
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test ($\chi^2_{(1)}$)	24.82		36.57		1.66	
p-value	0.000		0.000		0.198	
J Test ($\chi^2_{(3)}$)	2.292	2.155	2.563	2.452	1.486	1.472
p-value	0.514	0.541	0.464	0.484	0.686	0.689
# Obs.	92,117	92,117	52,333	52,333	39,784	39,784
# Clusters	8,580	8,580	4,448	4,448	4,132	4,132

Table 1.7: IV Regressions - Capital Expenditures and Acquisitions, Quarterly Beliefs

This table presents second stage IV regressions of corporate investment on cash flow, lagged investment, Tobin's Q, and estimated beliefs about the economy. The second-stage regression equation is:

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{CF_{i,t}}{A_{i,t}} + \beta_3 Q_{i,t-1} + \beta_4 \pi_t + \beta_5 \pi_t^2 + \epsilon_{i,t}$$

where the dependent variable is capital expenditures scaled by assets, and π is the belief that the economy is in an expansionary state. The Wald test is a test of the restriction that $\beta_4 + \beta_5 = 0$; the restricted regression imposes this restriction. Cash flow and lagged investment are endogenous controls and dependent variables in unreported first-stage regressions. The set of first stage instruments includes cash flow lagged 2, 3, and 4 years, and investment lagged 3 and 4 years; F-statistics test the joint significance of the instruments. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. All regressions include firm fixed effects; robust standard errors, clustered by firm, are shown in parentheses. The χ^2 statistic with three degrees of freedom (J Test) tests the over-identifying restrictions.

	All Firms		Manufacturing		Non-Manufacturing	
	Unrestricted	Restricted	Unrestricted	Restricted	Unrestricted	Restricted
Cash Flow _t	0.279*** (0.040)	0.279*** (0.040)	0.213*** (0.041)	0.212*** (0.041)	0.327*** (0.079)	0.327*** (0.079)
Investment _{t-1}	0.411*** (0.074)	0.410*** (0.074)	0.238*** (0.080)	0.234*** (0.080)	0.569*** (0.125)	0.570*** (0.125)
Q	0.008*** (0.001)	0.008*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.011*** (0.001)	0.011*** (0.001)
π	-0.032** (0.014)		-0.058*** (0.016)		0.013 (0.025)	
π^2	0.029*** (0.011)		0.046*** (0.012)		-0.001 (0.019)	
$\pi \times (1 - \pi)$		-0.022*** (0.005)		-0.019*** (0.006)		-0.028*** (0.010)
CF Stage 1 F-stat	114.24	114.14	81.11	81.06	41.06	41.02
p-value	0.000	0.000	0.000	0.000	0.000	0.000
I Stage 1 F-stat	106.97	106.76	66.57	66.57	45.95	45.79
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test ($\chi^2_{(1)}$)	0.64		6.81		3.04	
p-value	0.424		0.009		0.081	
J Test ($\chi^2_{(3)}$)	2.220	2.218	2.512	2.545	1.559	1.561
p-value	0.528	0.529	0.473	0.467	0.669	0.668
# Obs.	92,117	92,117	52,333	52,333	39,784	39,784
# Clusters	8,580	8,580	4,448	4,448	4,132	4,132

Chapter 2

An Empirical Analysis of Investment Over the Political Cycle

2.1 Introduction

Controlling for investment opportunities, the investment behavior of U.S. firms varies systematically over the political cycle. Consider the simple regression

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 Q_{i,t-1} + \beta_2 Election_t + \beta_3 (Q_{i,t-1} \times Election_t) + \epsilon_{i,t} \quad (2.1)$$

where I is corporate investment (capital expenditures plus acquisitions), A is book value of assets, Q is the market-to-book ratio, a proxy for marginal Q , and $Election$ is a dummy variable equal to one during Presidential election years and zero otherwise. For U.S. firms, the estimate of β_2 is negative: Across all firms from 1964 to 2008, the point estimate is -0.0021, with a t-statistic of -2.65. In economic terms, this amounts to a 2.43% reduction in investment expenditures during election years, holding investment opportunities constant.¹

Investment rates also differ systematically across periods of divided and single-party government. Investment tends to fall during periods of divided government, though this relationship varies with Q . For firms with low Q (a sign of few investment opportunities), investment decreases when government is divided; assuming a value of $Q = 1$, the effect is approximately a 4.03% reduction in investment expenditures as a percentage of assets. For firms with high Q (a sign of valuable investment opportunities), the effect turns positive; at the sample average value of $Q = 2.11$, the

¹The estimate of β_3 is -0.0004, and the average investment rate I/A is 0.1178. Using the sample average 2.1197 for Q , the calculation is $\% \Delta(I/A) = (-.0021 - .0004 \times 2.1197) / .1178 = -.0243$. Table 2.3 reports the regression estimates.

effect is a 0.79% increase in investment during periods of divided government. This dynamic effect is not very robust, however.²

Neoclassical investment theory has little to say about direct links between investment and the political environment. Indeed, policy should play a role only indirectly if investment opportunities vary with the political cycle. Loosely speaking, equation (2.1) captures this indirect channel through β_3 and the interaction term; β_2 and the dummy variable captures direct links.

What direct channel might be driving this result? From an empirical standpoint, the regression results provide evidence that political uncertainty and political outcomes play a role in corporate investment decisions. Indeed, political uncertainty should matter for a firm's investment decisions if regulatory policy affects either project cash flows or the (sunk) cost of an investment. Guided by these results, I study the effect of regulatory and public policy on corporate decisions. In particular, how does regulatory and political uncertainty affect a firm's decision to allocate capital to irreversible investment projects? Firms may find it optimal to forgo NPV-positive projects while retaining their right to invest at a later date, if delaying allows them to learn about the regulatory environment.

I study the link between policy uncertainty and investment through a straightforward re-thinking of the model presented in Chapter 1. In the theoretical framework, firms do not have perfect information about the regulatory environment and must form a belief about whether the environment is favorable or unfavorable. The regulatory environment governs the growth rate of cash flows, which is higher during periods of favorable policy. Growth rates are also unobservable, and beliefs inform the firm's estimates of growth rates and the value of a perpetual stream of cash flows.

Firms are also able to choose the scale of their investments. In addition to choosing the optimal time to invest, firms must choose the optimal amount of capital to invest. When beliefs are optimistic, so that estimated growth rates and present values are high, firms will want to commit a large amount of capital to take advantage of high growth. When beliefs are pessimistic, however, firms commit less capital since valuations are lower and projects appear less attractive.

Under this setup, the theoretical framework generates two key predictions:

1. Firms will invest when they have precise beliefs about the regulatory environment, regardless of whether beliefs are optimistic or pessimistic, but will delay when uncertainty is highest.
2. The incentive to delay is strongest when uncertainty resolves quickly.

The first prediction is a consequence of a firm's ability to choose the scale of its projects. The amount of capital that maximizes NPV differs in the two regulatory

²Simple regressions also suggest that investment rates are systematically lower during Republican presidential administrations. However, this effect does not remain after controlling for the economic and fiscal environment.

environments. When investment is irreversible, firms must guard against allocating an inefficient amount of capital to a project. Therefore, they may optimally delay projects to gather more information about regulatory policy. This result is intuitive, but it depends critically on allowing the scale of the project to vary. In standard real option models, where firms face just a timing problem, firms will only invest if uncertainty resolves in the favorable direction. When scale is a choice variable, firms invest even when uncertainty resolves unfavorably; they just choose smaller projects.

The second prediction is a consequence of learning. Forgoing an NPV-positive project can be optimal if delaying investment allows the firm to learn about the regulatory environment. The benefit can be substantial if delaying allows the firm to make a more informed capital allocation choice. When uncertainty resolves slowly, however, there is little to learn. In this case, the opportunity cost of delaying investment will outweigh the benefits gained from learning.

The election year results are consistent with the theoretical framework if 1) regulatory uncertainty is higher during election years, and 2) uncertainty resolves fairly quickly after an election as policy-makers take action. Empirically, I attempt to rule out other interpretations that are not consistent with the model. For example, a simple crowding-out story may drive this result. Politicians want to win reelection and, during election years, have incentives to spend more money in an effort to secure votes. If deficit spending draws resources away from private investment, then we would expect to see investment rates fall, all else equal. I control for potential crowding-out effects by including the federal budget deficit as an explanatory variable in the empirical analysis.

The crowding-out hypothesis highlights the central empirical challenge in taking the model to the data - political uncertainty is unobservable. The Presidential election cycle may capture the evolution of political uncertainty, but it may also proxy for other effects, such as politicians concerned with reelection. To minimize the possibility that the empirical results are spurious, I also include cash flows and lagged investment as firm-level controls and GDP growth as a macroeconomic control.

The empirical and theoretical frameworks build on previous work by Julio and Yook (2009) and Hassett and Metcalf (1999). Empirically, using ordinary least squares regressions, Julio and Yook find that the election year effect exists internationally as well. Across a sample of 48 countries, they find that during the year leading up to a national election (President or Prime Minister, as opposed to legislative elections), firms reduce investment expenditures by 5.3% on average. In addition to including an indicator variable for divided government, my analysis differs from theirs for two key reasons. First, because cash flows are endogenous to the corporate investment decision, the results of Julio and Yook may be due to the use of an inconsistent estimator. Indeed, I show that, for a specific non-manufacturing sub-sample of firms, least squares regressions easily reject the null hypothesis that investment is unrelated to the political cycle. However, after controlling for endogenous cash flows and potential autocorrelation in the error term, investment for this sample of firms appears to

have no relationship with political cycle, outside of effects associated with the fiscal environment.³

Second, my analysis suggests that the relationship between investment and the election cycle is much more dynamic. In particular, it strongly depends on the fiscal environment. In a balanced budget environment, investment indeed declines during election years, all else constant. However, firms react differently to federal budget deficits in election years so that when the country is running large deficits, investment can actually increase during election years. Failure to control for the fiscal environment may mask important dynamics in the relationship between investment and policy uncertainty.

In Hassett and Metcalf, the policy environment affects the value of a single project over time. The enactment and reversal of investment tax credits, which decrease the price of capital and increase the attractiveness of investment opportunities, arrive randomly and at discrete intervals. Similarly, Chen and Funke (2003) develop an incremental investment model where regulatory policy affects the profitability of the firm. Both of these papers assume that 1) the regulatory environment is known, and 2) policy change follows a Poisson arrival process. These assumptions are convenient in that they allow for closed-form solutions and comparative statics. Finally, Rodrik (1991) considers how potential reversals of policy reforms designed to stimulate investment may actually hinder investment; in his model, the probability of reversal is constant.

These previous models cannot capture changes in political uncertainty because of the memoryless property of the exponential distribution. Under a Poisson process, the expected time until the next policy change does not depend on the length of time since the last policy change; in these models, policy uncertainty is constant. To put it in the context of Bernanke (1983), beliefs about policy persistence never change. Incomplete information turns out to be an analytically convenient and economically intuitive way to introduce time-varying regulatory uncertainty.⁴

Nevertheless, incomplete information is a simplifying assumption. Intuitively, policy-makers may have preferences for regulations that favor either business or labor, and those preferences may not be immediately apparent during campaigns if a politician's incentives are to appeal to as many voters as possible. Furthermore, even if a politician's preferences are transparent, enactment of policy may be delegated to regulators within agencies who have their own set of preferences and are subject to industry capture. Spiller (1990) models the agency problems between the U.S. Congress and its regulators, both theoretically and empirically, and fails to reject the existence of an agency problem. While his framework suggests that Congress may use the budgeting process to discipline regulators, the empirical findings suggest that

³The discussion of the empirical specification in Section 2.3 provides more detail on the identification strategy and the associated econometric issues. Table 2.5 presents the regression results for this particular sub-sample.

⁴See Shreve (2004) for a detailed discussion of the memoryless property of Poisson processes.

Congressional control over regulatory agencies is far from perfect.

Finally, there is both anecdotal and empirical evidence that regulatory uncertainty is indeed a factor in firms' investment decisions. As reported in the journal *Refocus*, a 2005 survey by PricewaterhouseCoopers on the power supply industry found that “more than a third (39%) [of investors] say market reforms are damaging confidence, highlighting the dangers of inconsistent regulation, energy, tax, and environmental policies. The major survey of boardroom opinion inside utility companies also found that regulatory uncertainty is also affecting investment in renewables.”

2.2 Theoretical Review

I use the model presented in Chapter 1 as a tool to think about how policy uncertainty may affect investment. Because I discuss the model in full detail in the previous chapter, I provide a hands-off review of the relevant details here.

Model Setup

A firm has the option to invest K units of capital in a project. Investment is irreversible, and I assume that capital does not depreciate. The firm is risk-neutral, and its objective is to choose an investment time τ , as well as an investment level K , to maximize the expected net present value (NPV) of cash flows,

$$G = \max_{K, \tau} E_0 \left[\int_{\tau}^{\infty} e^{-ru} P(u) F(K_{\tau}) du - e^{-r\tau} K_{\tau} \right]. \quad (2.2)$$

After investment, the firm produces $F(K) = \log(1 + K)$ units of output in perpetuity. The output price is P , so that per-period cash flows are $P(t)F(K)$. The price of capital is normalized to unity.

I assume that the price of the output good is stochastic and follows a geometric process⁵

$$\frac{dP}{P} = \mu(\theta)dt + \sigma dZ_P \quad (2.3)$$

where P_Y is a standard Brownian motion and $\theta(t) \in \{\theta_f, \theta_u\}$ is a state variable describing the regulatory environment, corresponding to ‘favorable’ and ‘unfavorable.’ This is meant to capture periods of differing regulatory policy.

Shifts in θ are governed by a hidden two-state Markov process. I make the simplifying assumption that the probability of a change in policy is equally likely in both

⁵Assuming stochastic price is identical to assuming stochastic demand in Chapter 1 when the demand curve is perfectly elastic.

environments. Therefore, the transition matrix is

$$\mathbf{\Lambda} = \begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix}. \quad (2.4)$$

I make two assumptions about θ . The first assumption is that the regulatory environment affects the output price, and hence the cash flows of the investment project, through the growth rate. As an example, consider the pharmaceutical industry, where prices and cash flows may depend on whether regulators impose price controls on prescription drugs, allow less expensive drugs to be imported from abroad, or allow generic competition. A second example is the oil and gas industry, where regulators may explicitly focus on price as a mechanism to discourage carbon emissions, for example through tax policy.

Furthermore, note that ‘favorable’ need not imply the absence of regulation. In the case of farm subsidies or anti-dumping laws, regulations are often put in place to protect the cash flows of certain industries.⁶

I also assume the firm cannot directly observe the regulatory environment; only the regulator or policy maker knows his or her type. From a technical standpoint, this means that the firm does not observe either θ or dZ_P , nor do they observe $\mu(\theta)$; they only directly observe the output price P .

However, while not having complete information, the firm does see a noisy signal s about the regulatory environment,

$$ds = \theta dt + \eta \rho dZ_P + \eta \sqrt{1 - \rho^2} dZ_s. \quad (2.5)$$

Examples of noisy signals include regulatory decisions about other industries or products, public statements given by regulators or political candidates, and election results. The signal is noisy because regulatory policy is discretionary, changes in political leadership do not necessarily mean a change in regulatory and public policy, and campaigning is different from governing.

The Filtering Equations

Define the conditional expectations,

$$\begin{aligned} \hat{\theta} &= \pi \theta_f + (1 - \pi) \theta_u \\ \hat{\mu} &= \pi \mu_f + (1 - \pi) \mu_u, \end{aligned}$$

⁶I will typically assume that $\mu(\theta_f) > \mu(\theta_u)$.

the constants

$$\omega_P = \frac{\mu_f - \mu_u}{\sigma}$$

$$\omega_s = \frac{\theta_f - \theta_u}{\eta\sqrt{1 - \rho^2}} - \frac{\mu_f - \mu_u}{\sigma} \frac{\rho}{\sqrt{1 - \rho^2}},$$

and the adapted Brownian motions

$$d\hat{Z}_P = \frac{\mu(\theta) - \hat{\mu}}{\sigma} dt + dZ_Y$$

$$d\hat{Z}_s = \left(\frac{\theta - \hat{\theta}}{\eta\sqrt{1 - \rho^2}} - \frac{\mu(\theta) - \hat{\mu}}{\sigma} \frac{\rho}{\sqrt{1 - \rho^2}} \right) dt + dZ_s.$$

Then using the filtering results in Chapter 1 (or, alternatively, in Appendix A), the observed stochastic processes for price, signals, and posterior beliefs are

$$\frac{dP}{P} = \hat{\mu} dt + \sigma d\hat{Z}_P \quad (2.9a)$$

$$ds = \hat{\theta} dt + \eta \rho d\hat{Z}_P + \eta \sqrt{1 - \rho^2} d\hat{Z}_s \quad (2.9b)$$

$$d\pi = \lambda(1 - 2\pi) dt + \pi(1 - \pi) \left(\omega_P d\hat{Z}_P + \omega_s d\hat{Z}_s \right), \quad (2.9c)$$

In this simpler setting, because the probability of policy change is identical for both policy environments, beliefs revert to a mean of $\pi = 1/2$ at rate 2λ .

Project Valuation

Once the firm exercises its option to invest, the project generates per-period cash flows of $P(t)F(K)$ in perpetuity. Firms are risk-neutral and discount cash flows at a constant rate r . Using the results in Chapter 1, specifically equations (1.23) and (1.23), the discounted value of future cash flows is given by

$$V(P, \pi; K, t) = [A\pi + B(1 - \pi)] P(t)F(K) \quad (2.10)$$

with

$$A = \frac{r - \mu_u + 2\lambda}{(r - \mu_f)(r - \mu_u) + [\lambda(r - \mu_f) + \lambda(r - \mu_u)]}$$

$$B = \frac{r - \mu_f + 2\lambda}{(r - \mu_f)(r - \mu_u) + [\lambda(r - \mu_f) + \lambda(r - \mu_u)]}.$$

Finally, the value of the investment option must satisfy the partial differential

equation

$$\begin{aligned}
 rG = & [\pi\mu_f + (1 - \pi)\mu_u] PG_P + \frac{1}{2}\sigma^2 P^2 G_{PP} + \lambda(1 - 2\pi)G_\pi \\
 & + \frac{1}{2}\pi^2(1 - \pi)^2 (\omega_P^2 + \omega_s^2) G_{\pi\pi} + \pi(1 - \pi) (\mu_f - \mu_u) PG_{P\pi}.
 \end{aligned} \tag{2.12}$$

subject to value-matching and smooth-pasting conditions. When firms are able to choose the amount of capital allocated to a project, the capital decision is

$$K^* = \max \{ [A\pi + B(1 - \pi)] P^* - 1, 0 \}.$$

I refer readers to the numerical analysis presented in Chapter 1 for a detailed review of the properties of the solution. Briefly, Figures 1.4 and 1.6 present two key results. The first is that investment thresholds are increasing with uncertainty when firms have flexibility over both the timing and scale of their investment projects. The intuition is that, by making an investment decision during periods of high policy uncertainty, firms risk irreversibly allocating a suboptimal level of capital to a project. Therefore, prices need to be sufficiently ‘high’ before an investment opportunity becomes attractive. Contrast this to the case where capital is inflexible. In this case firms face only a timing decision - because irreversibly committing an inefficient level of capital is never a possibility, investing is always a dominant strategy as long as prices are higher than the $\pi = 0$ threshold.

The second result is that the timing of uncertainty resolution is important for determining investment. When uncertainty resolves quickly, which is likely to be the case for uncertainty associated with political elections, the benefits to delaying during periods of high uncertainty can be substantial. However, when uncertainty resolves slowly, or not at all, there is little benefit to delaying.

To summarize, the investment model predicts that firms will delay investment during periods of high policy uncertainty if 1) firms have flexibility over both the timing and scale of their investments, and 2) policy uncertainty resolves quickly.

2.3 Empirical Framework

2.3.1 Empirical Specification

To study the empirical relationship between political cycles and corporate investment, I use equation (2.1) as a starting point. To estimate the effect of policy uncertainty on investment, I run the following regression:

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 Q_{i,t-1} + \beta_2 U_t + W_{it}\theta + X_{it}\gamma + Y_t\delta + \epsilon_{i,t} \tag{2.13}$$

where I is corporate investment, A is the book value of assets, Q is a proxy for marginal Q , and U is a measure of political or regulatory uncertainty. Additionally, W is a set of endogenous firm-specific controls, X is a set of exogenous firm-specific controls, and Y is a set of exogenous macroeconomic controls. Under the null hypothesis, Q is a sufficient statistic for investment. Thus, under the null, only β_1 is nonzero.

Equation (2.13) presents several empirical challenges for testing. The first is that policy uncertainty is unobservable; testing (2.13) requires a good proxy. The statistical issue is that the proxy variable may capture effects unrelated to policy uncertainty, generating a spurious relationship. Essentially, there may be firm-level or macroeconomic variables in the error term $\epsilon_{i,t}$ that are correlated with the proxy for policy uncertainty. If so, this will bias the estimate of β_2 and potentially suggest a relationship that does not exist. Therefore, I include firm-level and macroeconomic controls W_{it} , X_{it} , and Y_t .

I consider two measures of policy uncertainty: an election year indicator equal to one during Presidential election years, and a divided government indicator equal to one during periods of divided government. Consider first the election year variable. If the outcome of an upcoming election is in doubt, uncertainty about regulatory and public policy is likely higher, to the extent that policy platforms differ across parties. Therefore, all else equal, the election year indicator should capture this effect, relative to non-election years.

Next, consider the divided government variable, where I define divided government as any scenario in which no single party has control simultaneously over the White House and both branches of Congress. This variable is meant to capture ideological variance and uncertainty about the legislative agenda - if the outcome of legislation is more uncertain during periods of divided government, then all else equal, the divided government indicator should capture this effect.

I also include an indicator for periods when the President is a member of the Republican party. I use this variable not to measure uncertainty, but rather to capture different policy environments. As the modeling framework suggests, regulatory uncertainty is important only if policy affects the cash flows generated by a project. Therefore, the Republican variable becomes economically important in the first stage cash flow regressions.

The exogenous firm-level controls include lagged investment and Q interacted with the political indicators. I use cash flows as an endogenous control:

$$W_{it} = \frac{CF_{i,t}}{A_{i,t}}, \quad X_{it} = \left[\frac{I_{i,t-1}}{A_{i,t-1}}, \text{ Interaction Terms} \right]. \quad (2.14)$$

I include investment as an exogenous control mainly for econometric reasons. If investment is persistent, shocks to investment will generate serially-correlated error terms. Including lagged investment is a straightforward way to generate error terms

with more desirable properties in the time series. (Blundell et al. (1992) and Brown and Petersen (2009).) Additionally, the political business cycle literature suggests that investment opportunities may vary with the political cycle for reasons that are unrelated to uncertainty. The interaction terms control for this possibility. If political cycles in investment are solely due to systematic variation of Q with the political cycle, coefficients on the interaction terms should be non-zero, and coefficients on the dummy variables should be zero.

A large body of literature, beginning with Fazzari et al. (1988), suggests that cash flow is an important determinant of investment, so I include it as an endogenous control. Although these studies typically argue that investment-cash flow sensitivities are evidence of financing constraints (but see Kaplan and Zingales (1997) for a counter-argument), this is only true if marginal Q is measured without error, a point made by Erickson and Whited (2000) and Lewellen and Lewellen (2010). An alternative rationale for including cash flow as a regressor is that cash flow may be a better proxy for marginal Q than the observed average Q . Indeed, in the model of Abel and Eberly (1997), where firms face convex adjustment costs, cash flows are perfectly correlated with marginal Q , while average Q is a biased measure of marginal Q .⁷

The macroeconomic controls are the growth rate of real gross domestic product, the federal budget deficit, and the deficit interacted with the election year dummy,

$$Y_t = [\% \Delta GDP, \text{ Deficit}, \text{ Deficit} \times \text{ Election}]. \quad (2.15)$$

Including GDP growth controls for business cycle effects and possible covariance of GDP with the political variables. I also include the control for government expenditures to help distinguish between two potential, but very different, interpretations of the negative relationship between investment and national elections. Under one interpretation, firms reduce investment expenditures during election years because of Bernanke's (1983) "bad news" principle: If uncertain election outcomes increase the perceived probability of negative regulatory change, firms may delay investment until some or all of the uncertainty is resolved. Alternatively, politicians motivated by reelection concerns may be more likely to spend money in election years in an effort to secure votes. If this increased spending is financed through a budget deficit, then government borrowing may crowd out private investment.

To disentangle these two effects, I include a control for the deficit, as well as an interaction term between the deficit and the election year dummy. If politicians' incentives are different during election years vs. non-election years, so that governments spend money for different reasons in election years, then we might expect corporate investment to respond differently to deficit spending in election years relative to non-

⁷Additionally, Novy-Marx (2006) argues from a theoretical standpoint that observed investment-cash flow sensitivities arise endogenously in general equilibrium due to the nonlinear relationship between demand and investment, Q , and cash flow.

election years. The interaction term captures this differential effect.

The second empirical challenge involves the inclusion of lagged investment in the vector of exogenous controls X_{it} , which is correlated with the unobserved firm fixed effect α_i . Failure to correct for unobserved heterogeneity will result in biased coefficients due to omitted variable bias. Two methods are available to deal with the fixed effect. The more common method, a within-groups transformation of the data, eliminates the fixed effect but in the process generates correlation between the transformed lagged dependent variable and the transformed error term. Unless a suitable instrument is available, the within-groups fixed effects estimator is inconsistent.

An alternative method, the first-differences fixed effects estimator, also generates correlation between the differenced lagged dependent variable and the differenced error term. However, as long as the error term is serially uncorrelated, deeper lags of investment (two periods and beyond), which are not available as instruments in the within-groups case, will be uncorrelated with the error term and make suitable instruments. This approach is due to Anderson and Hsiao (1982) and Arrelano and Bond (1991).

The final empirical consideration in estimation is that the endogenous control, cash flows, is simultaneously determined with investment. Therefore, the cash flow variable is likely correlated with the error term in the regression equation due to unmodeled operating decisions within the firm. Following Brown and Petersen (2009), I use lagged values of cash flow as instruments, as in the lagged investment case. While not necessarily economically meaningful, these are valid instruments if the error term in the regression is i.i.d. Intuitively, since lagged values are predetermined, they will be uncorrelated with the error term at time t . With more instruments than endogenous regressors, the system is over-identified; the χ^2 test of over-identifying restrictions is a test of orthogonality conditions.

2.3.2 Data Description

The empirical analysis in this chapter utilizes the same data on US firms as the analysis in Chapter 1, and I refer readers to the description in Chapter 1 for detail on the construction of the corporate data set. The only difference is that I also remove firms with missing values for cash and net working capital, whereas the Chapter 1 analysis keeps these observations.⁸

Briefly, firm investment and cash flow data come from the COMPUSTAT database; I use 1964 as the beginning date due to limited availability prior to 1964. I consider two measures of investment. The first, capital expenditures, is taken from the statement of cash flows. The second measure adds cash spent on acquisitions, also taken from the statement of cash flows. To measure cash flow, I use operating income before depreciation and amortization, taken from the income statement. Because

⁸The results are virtually identical when these observations are kept.

cash flow and investment are flow variables, I scale yearly values by the average of beginning-year and ending-year book value of assets.

To proxy for marginal Q , which is generally unobservable, I estimate average Q as the sum of market value of equity and book value of liabilities, all scaled by book value of assets. Average Q is an imperfect measure of marginal Q , which neoclassical investment theory identifies as the appropriate measure of the value of capital relative to replacement cost. Nevertheless, this proxy for average Q is standard in the literature.

In order to be included, firms must have non-missing values for Q , assets, investment, and cash flow. Additionally, I drop any firms with negative book value of equity (a sign of severe financial distress, which constrains investment), and those with market value of equity equal to zero. Furthermore, as previously mentioned, I remove observations with missing values for cash and net working capital. Finally, I remove any firms with headquarters outside the United States, or with financial statements reported in a currency other than the U.S. dollar. Given the size and influence of the U.S. economy, it is reasonable to believe that U.S. political considerations may influence international firms as well. However, preliminary tests suggested that removing these firms does not materially change the results.

Filtering down to non-missing observations and U.S.-based firms results in an unbalanced panel of 155,078 firm-year observations over the sample. As a final adjustment, because of concerns about measurement error and data coding, I Winsorize investment, cash flow, and Q at the 1st and 99th percentiles. Table 2.2 presents summary statistics on these variables before and after the adjustment.

For the macroeconomic controls, I collect data on GDP, government expenditures, and taxes from the FRED database at the St. Louis Federal Reserve Bank. Table 2.1 provides the data codes and descriptions. I calculate the budget deficit as government expenditures less government tax receipts, scaled by government expenditures, so that deficits are represented by positive values and surpluses by negative values. Although the COMPUSTAT and fiscal data are reported in nominal terms, I use real instead of nominal GDP as a macroeconomic control since scaling tends to remove price level effects from the time series.

2.3.3 Estimation

Estimation of the empirical specifications is through first-differenced two-stage least squares regression on an unbalanced panel. I estimate first stage regressions for both cash flow and lagged investment, and the excluded instrument set includes cash flow lagged 2, 3, and 4 years, and investment lagged 3 and 4 years. Following Stock and Watson (2008), all regressions include firm fixed effects with standard errors clustered by firm.

The first stage estimation of cash flow serves two purposes. First, as already mentioned, cash flow is endogenous to the investment decision; failure to control for

endogenous regressors may result in inconsistent estimates. Second, in order for policy uncertainty to matter to the investment decision, cash flow must vary systematically across policy regimes. Using the Republican presidential indicator as a proxy for the policy environment, I can test this condition directly.

Nevertheless, while instrumental variable techniques are appropriate, I first estimate (2.13) using least squares techniques. The purpose is twofold: First, in their international study of investment over political cycles, Julio and Yook (2009) use ordinary least squares with fixed effects. Comparing results to theirs provides a useful benchmark. Second, least squares regressions help to highlight the potential dangers of drawing inferences from inconsistent estimators.

Ordinary Least Squares Regressions

Table 2.4 presents benchmark ordinary least squares regressions of both measures of investment on cash flow, Q , GDP growth, and the presidential election indicator. This is the same specification studied by Julio and Yook (2009). In each case, I also break the sample into manufacturing and non-manufacturing firms, since irreversibility, which has its roots in asset specificity, is likely to be more important for manufacturing firms. In all specifications, the coefficient on the election year indicator is negative and statistically significant at the 1% level. Point estimates range from -0.002 for manufacturing firms when using capital expenditures as the measure of investment, to -0.007 for non-manufacturing firms when using the sum of capital expenditures and acquisitions as the measure of investment.

In terms of economic magnitudes, these numbers imply a reduction in investment rates on the order of 2.1% to 4.7% during election years, depending on the sample and the measure of investment. These results are qualitatively similar to Julio and Yook (2009), although the magnitudes are a bit smaller.

Table 2.5 explores the non-manufacturing sector a bit further. In this set of regressions, I estimate the effect of the presidential election cycle using the sum of capital expenditures and acquisitions as the measure of investment. The first three specifications are estimated using ordinary least squares with fixed effects, the standard econometric technique in the literature for estimating investment regressions, while the fourth specification is estimated using instrumental variables and the within-groups fixed effects estimator. Therefore, the first four specifications contain one or more explanatory variables that, for some reason or another, are correlated with the error term. Only the last specification, which uses the first-differences estimator, is consistent, and only the last specification fails to reject the null hypothesis that election years have no effect on investment policy. Note also that the χ^2 test of over-identifying restrictions easily rejects the within-groups instrumental variables model, while failing to reject the first-differences model. This suggests that empirical studies that rely on ordinary least squares estimates are not as reliable as thought.⁹

⁹Table 2.5 reports negative R^2 statistics for the instrumental variables specifications of the model.

Instrumental Variables Regressions

Turning attention now to instrumental variables regressions, Tables 2.6 and 2.7 present second-stage regression estimates of investment for all firms in the sample, as well as first-stage estimates of cash flow and lagged investment. Table 2.6 uses capital expenditures as the measure of investment, while Table 2.7 uses capital expenditures plus cash spent on acquisitions as the measure of investment. These two tables reveal several key results.

First, based on first-stage regressions, cash flows are systematically lower during Republican presidential administrations, even after controlling for GDP growth. In both cases, the coefficient is significant at the 1% level. This result is important since a necessary assumption of the investment model is that prices and cash flows differ across regulatory environments. Furthermore, F-statistics fail to reject the hypothesis of joint insignificance of the excluded lagged instruments.

Second, χ^2 tests of the over-identifying restrictions fail to reject the orthogonality conditions, although p-values are much higher when the second measure of investment is used. Indeed, the test almost rejects the model using only capital expenditures. Nevertheless, the first-stage F tests combined with the χ^2 tests give validity to the excluded instrument set and estimation procedure.

Finally, the point estimates on the election year and divided government indicators are negative and statistically significant at at least the 5% level, and the 1% level when acquisitions are included, which is consistent with the results of Julio and Yook (2009). Using the results in Table 2.7, the partial effect of divided government on investment is given by

$$\frac{\partial I/A}{\partial \text{Divided}} = -0.006 + 0.002 \times Q,$$

so that investment declines during periods of divided government unless Q is sufficiently high ($Q > 3.42$ to be exact). In the sample, the average investment rate is 0.118, while the average value of Q is 2.120. This translates into a reduction in investment rates of approximately 2% during periods of divided government, when using both capital expenditures and acquisitions as the measure of investment. The effect is the same when using only capital expenditures, and ranges from about -1% for manufacturing firms to -3% for non-manufacturing firms. During periods of divided government, ideological variance is higher. To the extent that policy uncertainty is associated with ideological variance, then the relationship between investment and divided government is consistent with a model where firms decrease investment in

This is always a possibility with IV regressions because the residual sum of squares is calculated using actual values of endogenous and exogenous regressors, while the total sum of squares is calculated using regressors projected onto the set of instruments. As a result, there is no guarantee that the R^2 statistic will lie between zero and one, nor R^2 even a meaningful statistic. For this reason, I do not report this statistic in the first differences regressions.

response to higher policy uncertainty.

Coefficients on the election year indicator also suggest that investment rates decline during election years. However, the coefficients on the deficit control and deficit-election interaction term reveal a relationship between investment and election cycles that is much more dynamic.

To be specific, in a balanced budget environment, firms indeed reduce investment expenditures during election years. In Table 2.7, the coefficient on the election year indicator is -0.005, which translates into an approximately 4.7% decrease in investment during election years. However, the regressions indicate that firms' reactions to budget deficits are markedly different during election years. The coefficient on the deficit control is -0.069, implying that firms typically reduce investment expenditures as deficits increase. This is the often-mentioned 'crowding-out' effect, where the need to finance deficits draws resources away from private investment. However, during election years, this effect almost completely disappears: The coefficient on the interaction term is 0.061, suggesting that the overall response to deficits is only -0.008 during election years.

This result may not be surprising if deficit spending is different during election years relative to non-election years. If election-year deficit spending is geared more towards stimulative projects, (e.g. investment tax credits), then it is reasonable to expect that firms may react to deficits differently during election years. However, regardless of the interpretation, the upshot is that the elimination of deficit effects means that investment can actually increase in election years. To see this, note that the partial effect of election years on investment is

$$\frac{\partial I/A}{\partial \text{Election}} = -0.005 + 0.061 \times \text{Deficit}.$$

which is positive for any budget deficit greater than 9% of government expenditures. The U.S. government typically runs a deficit, and the average value of the deficit in the sample is 10.1% of expenditures. Thus, on average, investment rates actually increase during Presidential election years, and we can break the total effect into a deficit effect, which is positive, and an additional effect, which is negative. It is this additional effect which I associate with policy uncertainty. While other interpretations may exist, the coefficients on the election year indicator are consistent with the hypothesis that increased regulatory and policy uncertainty associated with political elections has a dampening effect on overall investment.

2.4 Concluding Remarks

Uncertainty plays a critical role in determining when and how much capital firms allocate to irreversible investment projects. While price and cash flow risk are perhaps the most fundamental uncertainties, there are numerous risks the firm faces and must

consider in decision-making.

One such risk is regulatory and policy uncertainty. The empirical results presented suggest that firms alter investment policy in response to changing political environments and regulatory uncertainty. Firms care about political uncertainty if regulatory and public policy affects the cash flows of their investments.

The theoretical analysis developed in Chapter 1 and adapted to a policy environment provides a real option interpretation of the empirical results: Firms may find it optimal to forgo NPV positive projects if delaying allows them to gather more information about the regulatory environment. Of course, the model is not necessarily one of regulatory uncertainty per se; for example, it cannot distinguish between regulatory uncertainty and business cycle uncertainty. However, the usefulness of the model is in its descriptive properties because it provides a mechanism for the effect of policy uncertainty on investment that is consistent with the data: When uncertainty resolves quickly and firms have flexibility over the scale of their investments, political uncertainty can negatively affect corporate investment.

2.5 Tables

Table 2.1: **Description of FRED variables**

Federal Reserve Economic Data, from the Federal Reserve Bank of St. Louis.
<http://research.stlouisfed.org/fred2/>

Data Series	Descriptor
GDPCA	Real Gross Domestic Product
AFEXPND	Federal Government: Current Expenditures
AFRECPT	Federal Government: Current Receipts

Table 2.2: Summary Statistics - Investment, Cash Flow, and Tobin's Q

This table presents summary statistics for corporate financial data from the COMPUSTAT database, before and after adjusting for outliers. Capital expenditures, acquisitions, and cash flow are scaled by the average of beginning and ending book value of assets, and Tobin's Q is calculated as the market value of equity plus debt, scaled by book assets. To adjust for outliers, variables are Winsorized at the 1st and 99th percentiles.

Panel A: Unadjusted				
	Capital Expenditures	CapEx + Acquisitions	Cash Flow	Tobin's Q
Mean	0.097	0.120	0.106	2.458
Median	0.067	0.080	0.157	1.336
Standard Deviation	0.116	0.145	0.412	20.499
Skewness	19.439	11.249	39.874	230.248
Kurtosis	2,120.872	864.501	4,120.480	68,214.270
Minimum	-0.209	-1.526	55.628	0.003
Maximum	15.569	15.569	3.670	6,530.038
1st Percentile	0.001	0.000	1.116	0.421
99th Percentile	0.525	0.692	0.599	15.709
Panel B: Adjusted				
	Capital Expenditures	CapEx + Acquisitions	Cash Flow	Tobin's Q
Mean	0.094	0.118	0.115	2.120
Median	0.067	0.080	0.157	1.336
Standard Deviation	0.094	0.123	0.254	2.335
Skewness	2.212	2.305	-2.233	3.545
Kurtosis	8.932	9.345	10.422	17.907
Minimum	0.001	0.000	-1.116	0.421
Maximum	0.525	0.692	0.599	15.709
1st Percentile	0.001	0.000	-1.116	0.421
99th Percentile	0.525	0.692	0.599	15.709

Table 2.3: **OLS Regressions - Investment and Political Cycles**

This table presents OLS regressions of corporate investment on Tobin's Q and political cycle indicators. The regression equation is:

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 Q_{i,t-1} + \beta_2 Political_t + \beta_3 (Q_{i,t-1} \times Political_t) + \epsilon_{i,t}$$

where the dependent variable is the sum of capital expenditures and acquisitions, scaled by assets. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. All regressions include firm fixed effects; robust standard errors, clustered by firm, are shown in parentheses.

	(1)	(2)	(3)
Q	0.011*** (0.000)	0.010*** (0.000)	0.010*** (0.000)
Election	-0.002*** (0.001)		
Election×Q	0.000 (0.000)		
Divided Government		-0.002** (0.001)	
Divided×Q		0.002*** (0.000)	
Republican President			-0.016*** (0.001)
Republican×Q			0.002*** (0.000)
Constant	0.095*** (0.001)	0.097*** (0.001)	0.105*** (0.001)
R ²	0.031	0.031	0.034
# Obs.	155,078	155,078	155,078
# Clusters	14,838	14,838	14,838

Table 2.4: **Least Squares Regressions**

This table presents OLS regressions of corporate investment on cash flow, Tobin's Q, real GDP growth, and an election year indicator variable. The regression equation is:

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 \frac{CF_{i,t}}{A_{i,t}} + \beta_2 Q_{i,t-1} + \beta_3 \Delta GDP_t + \beta_4 Election_t + \epsilon_{i,t}$$

where the dependent variable is either capital expenditures or the sum of capital expenditures and acquisitions, scaled by assets. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. All specifications include firm fixed effects; robust standard errors, clustered by firm, are shown in parentheses.

	Capital Expenditures			Cap. Ex. plus Acquisitions		
	All Firms	Manuf.	Non-Manuf.	All Firms	Manuf.	Non-Manuf.
Cash Flow _t	0.053*** (0.002)	0.051*** (0.003)	0.055*** (0.004)	0.061*** (0.003)	0.058*** (0.003)	0.066*** (0.005)
Q	0.008*** (0.000)	0.007*** (0.000)	0.010*** (0.000)	0.010*** (0.000)	0.008*** (0.000)	0.013*** (0.000)
ΔGDP _t	0.062*** (0.011)	0.095*** (0.013)	0.020 (0.020)	0.153*** (0.015)	0.175*** (0.017)	0.126*** (0.025)
Election	-0.003*** (0.000)	-0.002*** (0.000)	-0.004*** (0.001)	-0.005*** (0.001)	-0.003*** (0.001)	-0.007*** (0.001)
R ²	0.052	0.054	0.053	0.040	0.035	0.047
Average Investment	0.094	0.079	0.113	0.118	0.099	0.140
Average Q	2.120	2.138	2.098	2.120	2.138	2.098
# Obs.	153,460	83,483	69,977	153,460	83,483	69,977
# Clusters	13,220	6,497	6,723	13,220	6,497	6,723

Table 2.5: **Least Squares vs. Instrumental Variables**

This table presents OLS and IV regressions of corporate investment on Tobin's Q , an election year indicator, and firm and macroeconomic controls. The regression equation is:

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{CF_{i,t}}{A_{i,t}} + \beta_3 Q_{i,t-1} + \beta_4 U_t + X_{it}\gamma + Y_t\delta + \epsilon_{i,t}$$

where the dependent variable is capital expenditures plus acquisitions scaled by assets and Q is Tobin's Q . U_t is a political cycle indicator, X is a set of exogenous firm controls, and Y is a set of exogenous macroeconomic controls. Cash flow and lagged investment are endogenous controls in the IV regressions; the set of first stage instruments includes cash flow lagged 2, 3, and 4 years, and investment lagged 3 and 4 years. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. All specifications include firm fixed effects; 'FE' stands for the within-groups fixed effects estimator, while 'FD' stands for the first differences estimator. Robust standard errors, clustered by firm, are shown in parentheses. The χ^2 statistic with three degrees of freedom tests the over-identifying restrictions.

	OLS-FE	OLS-FE	OLS-FE	IV-FE	IV-FD
Cash Flow _{<i>t</i>}	0.066*** (0.005)	0.068*** (0.005)	0.068*** (0.005)	0.354*** (0.081)	0.338*** (0.077)
Investment _{<i>t-1</i>}		0.182*** (0.007)	0.181*** (0.007)	0.074 (0.121)	0.545*** (0.124)
Q	0.013*** (0.000)	0.011*** (0.001)	0.011*** (0.001)	0.009*** (0.001)	0.011*** (0.002)
Δ GDP _{<i>t</i>}	0.126*** (0.025)	0.152*** (0.023)	0.114*** (0.024)	0.004 (0.046)	0.056 (0.042)
Budget Deficit			-0.044*** (0.008)	-0.070*** (0.014)	-0.057*** (0.017)
Election Year	-0.007*** (0.001)	-0.005*** (0.001)	-0.008*** (0.002)	-0.009*** (0.003)	-0.005 (0.003)
Election×Deficit			0.034*** (0.012)	0.038*** (0.014)	0.037* (0.019)
Election×Q			0.000 (0.001)	0.001 (0.001)	0.000 (0.001)
R ²	0.047	0.077	0.078	-0.041	-0.793
χ^2				16.128	1.081
p-value				0.001	0.782
# Obs.	69,977	61,418	61,418	42,478	38,235
# Clusters	6,723	5,893	5,893	3,998	3,993

Table 2.8: **Capital Expenditures, Manufacturing Firms**

This table presents IV regressions of corporate investment on Tobin's Q, political cycle indicators, and firm and macroeconomic controls. The second stage regression is:

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{CF_{i,t}}{A_{i,t}} + \beta_3 Q_{i,t-1} + \beta_4 U_t + X_{it}\gamma + Y_t\delta + \epsilon_{i,t}$$

where the dependent variable is capital expenditures scaled by assets and Q is Tobin's Q. U_t is a set of political cycle indicators, X is a set of exogenous firm controls, and Y is a set of exogenous macroeconomic controls. Cash flow and lagged investment are endogenous controls and dependent variables in first-stage regressions. The set of first stage instruments includes cash flow lagged 2, 3, and 4 years, and investment lagged 3 and 4 years. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. All specifications include firm fixed effects; robust standard errors, clustered by firm, are shown in parentheses. The F-statistic tests the joint significance of the instruments, while the χ^2 statistic with three degrees of freedom tests the over-identifying restrictions.

	Manufacturing Firms			Non-Manufacturing Firms		
	CF _t	I _{t-1}	I _t	CF _t	I _{t-1}	I _t
Cash Flow _t			0.135*** (0.019)			0.180*** (0.036)
Investment _{t-1}			0.220*** (0.038)			0.372*** (0.069)
Q	0.007*** (0.002)	0.002*** (0.000)	0.002*** (0.001)	0.007*** (0.002)	0.002** (0.001)	0.003*** (0.001)
ΔGDP _t	0.249*** (0.022)	-0.175*** (0.013)	-0.040*** (0.015)	0.108*** (0.024)	-0.200*** (0.019)	-0.006 (0.026)
Budget Deficit	-0.102*** (0.010)	-0.023*** (0.005)	-0.064*** (0.006)	-0.047*** (0.012)	-0.029*** (0.008)	-0.071*** (0.009)
Election Year	-0.001 (0.002)	-0.002** (0.001)	-0.003*** (0.001)	0.002 (0.003)	-0.004** (0.002)	-0.002 (0.002)
Election×Deficit	0.013 (0.013)	0.007 (0.005)	0.042*** (0.006)	0.039 (0.014)	0.016* (0.009)	0.020* (0.010)
Election×Q	0.001 (0.001)	0.000 (0.000)	0.000 (0.000)	-0.004*** (0.001)	0.000 (0.001)	0.000 (0.001)
Divided Government	-0.002 (0.003)	0.006*** (0.001)	-0.002 (0.001)	0.004 (0.004)	0.007*** (0.002)	-0.009*** (0.002)
Divided×Q	0.002 (0.001)	0.000 (0.000)	0.001 (0.000)	-0.001 (0.002)	0.000 (0.001)	0.003*** (0.001)
Republican President	-0.006 (0.004)	-0.005*** (0.001)	-0.005*** (0.001)	-0.013*** (0.004)	-0.009*** (0.002)	0.002 (0.003)
Republican×Q	-0.003* (0.002)	0.000 (0.000)	0.001 (0.001)	0.002 (0.002)	0.002* (0.001)	-0.001 (0.001)
CF _{t-2}	-0.137*** (0.010)	0.054*** (0.003)		-0.153*** (0.014)	0.060*** (0.005)	
CF _{t-3}	-0.092*** (0.009)	0.010*** (0.002)		-0.087*** (0.012)	0.015*** (0.004)	
CF _{t-4}	-0.054*** (0.010)	0.004 (0.003)		-0.061*** (0.012)	0.007 (0.004)	
I _{t-3}	-0.012 (0.011)	-0.150*** (0.006)		0.011 (0.010)	-0.122*** (0.009)	
I _{t-4}	-0.014 (0.010)	-0.061*** (0.007)		-0.009 (0.009)	-0.067*** (0.008)	
F-stat	79.26	181.91		39.27	80.98	
p-value	0.000	0.000		0.000	0.000	
χ ²			8.40			1.454
p-value			0.038			0.693
# Obs.	51,646	51,646	51,646	38,235	38,235	38,235
# Clusters	4,429	4,429	4,429	3,993	3,993	3,993

Table 2.9: Capital Expenditures and Acquisitions, Manufacturing Firms

This table presents IV regressions of corporate investment on Tobin's Q, political cycle indicators, and firm and macroeconomic controls. The second stage regression is:

$$\frac{I_{i,t}}{A_{i,t}} = \alpha_i + \beta_1 \frac{I_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{CF_{i,t}}{A_{i,t}} + \beta_3 Q_{i,t-1} + \beta_4 U_t + X_{it}\gamma + Y_t\delta + \epsilon_{i,t}$$

where the dependent variable is capital expenditures scaled by assets and Q is Tobin's Q. U_t is a set of political cycle indicators, X is a set of exogenous firm controls, and Y is a set of exogenous macroeconomic controls. Cash flow and lagged investment are endogenous controls and dependent variables in first-stage regressions. The set of first stage instruments includes cash flow lagged 2, 3, and 4 years, and investment lagged 3 and 4 years. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. All specifications include firm fixed effects; robust standard errors, clustered by firm, are shown in parentheses. The F-statistic tests the joint significance of the instruments, while the χ^2 statistic with three degrees of freedom tests the over-identifying restrictions.

	Manufacturing Firms			Non-Manufacturing Firms		
	CF _t	I _{t-1}	I _t	CF _t	I _{t-1}	I _t
Cash Flow _t			0.208*** (0.041)			0.337*** (0.076)
Investment _{t-1}			0.231*** (0.080)			0.541*** (0.123)
Q	0.007*** (0.002)	0.000 (0.001)	0.005*** (0.001)	0.007*** (0.002)	-0.002 (0.001)	0.009*** (0.002)
ΔGDP _t	0.247*** (0.022)	-0.178*** (0.021)	-0.057** (0.025)	0.108*** (0.024)	-0.209*** (0.027)	0.051 (0.042)
Budget Deficit	-0.102*** (0.010)	-0.052*** (0.008)	-0.083*** (0.012)	-0.047*** (0.012)	-0.047*** (0.011)	-0.062*** (0.017)
Election Year	-0.001 (0.002)	-0.004* (0.002)	-0.006** (0.002)	0.002 (0.003)	-0.004* (0.002)	-0.005 (0.003)
Election×Deficit	0.013 (0.013)	0.008 (0.013)	0.076*** (0.015)	0.039*** (0.014)	0.004 (0.015)	0.042* (0.021)
Election×Q	0.001 (0.001)	-0.001 (0.000)	0.000 (0.001)	-0.004*** (0.001)	0.001 (0.001)	-0.001 (0.001)
Divided Government	-0.002 (0.003)	0.010*** (0.002)	-0.002 (0.003)	0.004 (0.004)	0.011*** (0.003)	-0.012** (0.004)
Divided×Q	0.002 (0.001)	0.000 (0.001)	0.000 (0.001)	-0.001 (0.002)	0.000 (0.001)	0.004** (0.002)
Republican President	-0.006 (0.004)	-0.013*** (0.003)	-0.007** (0.003)	-0.013*** (0.004)	-0.015*** (0.003)	0.003 (0.005)
Republican×Q	-0.003* (0.002)	0.000 (0.001)	0.001 (0.001)	0.002 (0.002)	0.002 (0.001)	-0.001 (0.002)
CF _{t-2}	-0.137*** (0.010)	0.083*** (0.005)		-0.153*** (0.014)	0.087*** (0.008)	
CF _{t-3}	-0.092*** (0.009)	0.013*** (0.004)		-0.086*** (0.012)	0.023*** (0.007)	
CF _{t-4}	-0.055*** (0.010)	0.003 (0.004)		-0.060*** (0.012)	0.006 (0.007)	
I _{t-3}	-0.004 (0.005)	-0.053*** (0.008)		0.006 (0.006)	-0.067*** (0.008)	
I _{t-4}	-0.006 (0.006)	-0.023*** (0.007)		0.005 (0.006)	-0.046*** (0.007)	
F-stat	79.27	64.51		38.96	47.89	
p-value	0.000	0.000		0.000	0.000	
χ ²			2.234			1.253
p-value			0.525			0.740
# Obs.	51,646	51,646	51,646	38,235	38,235	38,235
# Clusters	4,429	4,429	4,429	3,993	3,993	3,993

Chapter 3

Dynamic Currency Hedging

3.1 Introduction

What is the optimal currency hedge for a risk-averse investor? The conventional wisdom/advice is that investors should hedge away the currency risk of their international investments: Foreign assets are denominated in foreign currency, so domestic investors (who care about payoffs in domestic currency) must convert the foreign currency payoffs to domestic currency at uncertain exchange rates. In other words, a long position in, for example, foreign stock implicitly carries a long position in foreign currency as well. If the foreign currency risk premium is zero, exchange rate uncertainty increases the variance of an international investments without a corresponding increase in expected return; investors can reduce the volatility of their international portfolio with no loss in expected return through an offsetting short position in short-term foreign bonds.

Indeed, this is the conclusion reached by Solnik (1974). When the currency risk premium is zero, and exchange rate returns are uncorrelated with foreign asset returns, a complete hedge is optimal - every dollar invested long (short) in foreign assets is offset with a corresponding short (long) position in foreign currency, so that the net foreign currency position is zero.

Solnik's analysis can be viewed as a benchmark model, and there are reasons why a complete hedge may be suboptimal. For example, if a foreign currency tends to appreciate when equity values fall, then currency can serve as an effective hedge in declining equity environments, and a net long position may be optimal. In an empirical exercise, Campbell et al. (2008) study the effects of correlation in detail. They find that, over the 30-year period from 1975 to 2005, a risk-minimizing global investor would have benefitted from net long positions in the U.S. Dollar, the Euro, and the Swiss Franc, and short positions in other major currencies. Their interpretation is that the Dollar, Euro, and Swiss Franc are stores of value which tend to appreciate when asset values fall.

A second reason why a net zero position in foreign currency may be suboptimal

is predictability in speculative currency returns, studied by the large literature on the forward premium anomaly.¹ Briefly, the forward premium anomaly refers to the empirical finding that excess returns to currency speculation are predictable by the interest differential, whereas the theory of uncovered interest parity concludes that this predictability should not exist.

One interpretation of this result is a time-varying risk premium in foreign exchange. The portfolio choice implication is that investors can dynamically adjust their currency hedge position in response to the risk premium (or a proxy for the risk premium such as the interest differential) to increase their Sharpe ratios. Such strategies are studied in detail within a speculative framework by Burnside et al. (2008). Additionally, Della Corte et al. (2008) find that a dynamic speculative strategy that conditions on the forward premium outperforms a dynamic strategy that follows the random walk benchmark, and that an investor with quadratic utility will pay a high performance fee to switch to this strategy.²

Of course, if the risk premium is unobservable, the success of such a dynamic strategy depends on how well the predictor variable proxies for the risk premium. Indeed, recent research by Neely and Weller (2000) and Villanueva (2007) suggests that predictability is much weaker than implied by the in-sample evidence. Neely and Weller find that VAR-based estimates of the risk premium fail to outperform a constant risk premium in out-of-sample tests. Villanueva finds that excess return predictions based on interest parity regressions fail to beat forecasts based on uncovered interest parity, also out-of-sample. Both studies rely on mean-squared prediction errors as the appropriate measure of out-of-sample predictability.

Empirical evidence on the effectiveness of currency hedging is inconclusive. As mentioned above, Campbell et al. (2008) find that deviations from a complete hedge can reduce the total risk of a portfolio due to non-zero correlations. Along another dimension, Froot (1993) studies the effectiveness of hedging for long-horizon investors. He finds that, while complete hedging is effective at reducing portfolio risk over short horizons, the opposite may hold true at long horizons due to long-run mean reversion in exchange rates. Campbell, et al., however, do not find this effect in their data set.

In this paper I study the optimal dynamic currency hedging strategy for an investor who is concerned about the time-varying risk premium, as well as the associated uncertainty about the risk premium, which is unobservable. Investors must filter the risk premium from observable asset prices, and therefore face estimation risk. The estimated risk premium, which is time varying, introduces a speculative motive for holding foreign exchange, which in turn moves the optimal hedge away from a complete one. However, estimation risk has an offsetting effect, and considerably tempers the dynamic strategy of a myopic investor.

¹See Engel (1996) and Sarno (2005) for surveys of the forward premium anomaly literature.

²This result is termed an ‘anomaly’ because, to date, equilibrium models have generally failed to identify a preference-based risk premium that can explain the excess returns to currency speculation.

3.2 Currency Hedging with Constant Risk Premia

In this section I present a simple portfolio choice model to illustrate the effects of parameter uncertainty on optimal currency hedging. I model the problem of a domestic investor with access to a foreign bond and stock, as well as a domestic bond. To isolate the effects of estimation risk, all parameters are constant, and the correlation between the foreign stock and the exchange rate is zero. This is a partial equilibrium environment, and the economic agent described is not necessarily a representative agent.

3.2.1 Asset Market Dynamics

The investment opportunity set includes domestic and foreign risk-free bonds, and a foreign stock. For convenience, I omit a domestic stock from the investment portfolio. Although it would be straightforward to include this asset, I wish to focus mainly on the currency risk hedging problem of a domestic investor - in the simplified framework presented here, a domestic stock would alter the hedging decision only if domestic equity is correlated with the foreign exchange rate.

The foreign exchange rate X_t , defined as the domestic price of one unit of foreign currency (e.g. dollars per pound for a U.S.-based investor), obeys a stochastic differential equation,

$$\frac{dX_t}{X_t} = [r_d - r_f + \lambda_X] dt + \sigma_X dZ_{X_t}, \quad (3.1)$$

where Z_{X_t} is a standard Brownian motion. The drift (or expected depreciation) of the exchange rate is determined by the difference between domestic and foreign risk-free interest rates (assumed constant), $r_d - r_f$, and a constant risk premium, λ_X . The diffusion coefficient σ_X is constant and strictly positive. Both the risk premium and the Brownian motion governing shocks to the exchange rate are unobservable.

Although I model the exchange rate process as exogenous, the functional form of the drift arises endogenously in general equilibrium models. Holding the risk premium constant, the depreciation of the exchange rate is driven by cross-country differences in interest rates. Taking the U.S. as the domestic country, if the U.S. - U.K. interest differential is +2%, then an investor in British bonds would expect the pound to appreciate relative to the dollar (intuitively, a capital gain) to compensate for the lower interest rate (or income yield). The risk premium enters due to the assumption of convex preferences.³

The foreign stock S_t , denominated in foreign currency, follows a geometric Brownian motion process,

³See Backus et al. (2001) for a discussion of how (3.1) arises endogenously.

$$\frac{dS_t}{S_t} = [r_f + \lambda_S] dt + \sigma_S dZ_{St}, \quad (3.2)$$

where Z_{St} is a Brownian motion orthogonal to Z_{Xt} . The variance of the foreign stock, σ_S^2 , is constant and strictly positive, and the stock return is uncorrelated with the currency return. To maintain focus on the currency risk premium, I assume the foreign equity risk premium λ_S is constant and observable.

Domestic and foreign interest rates are constant; bond prices are denominated in local currency and evolve deterministically over time:

$$\frac{dB_{dt}}{B_{dt}} = r_d dt, \quad \frac{dB_{ft}}{B_{ft}} = r_f dt, \quad (3.3)$$

with initial conditions $B_{d0} = B_{f0} = 1$.⁴

3.2.2 The Investor's Learning Problem

Investors observe the foreign exchange rate X_t , as well as the instantaneous currency return dX_t/X_t . However, they do not observe either the risk premium λ_X or the Brownian motion Z_{Xt} , and face an inference problem. Investors use their observations of the exchange rate to form an estimate of the risk premium, which I call $\hat{\lambda}_{Xt}$, and update their estimate via Bayes's Rule as they observe new prices.

The Inference Process

The optimal filtering rules follow from Theorems 11.1 and 12.1 in Liptser and Shirayev (1978), hereafter LS. Denote the investor's optimal estimate of the risk premium (in the sense of minimizing mean-squared error) as $\hat{\lambda}_{Xt} = E[\lambda_X | \mathcal{I}_t]$, with variance $\nu_t = E[(\hat{\lambda}_{Xt} - \lambda_X)^2 | \mathcal{I}_t]$, and furthermore assume a Normal prior, i.e. $\hat{\lambda}_{Xt} \sim \mathcal{N}(\hat{\lambda}_{X0}, \nu_0)$. Then the conditional (posterior) distribution at each date $t \geq 0$ is also normal, with moments that satisfy,

$$d\hat{\lambda}_{Xt} = \frac{\nu(t)}{\sigma_X^2} \left[\frac{dX_t}{X_t} - (r_d - r_f + \hat{\lambda}_{Xt}) dt \right] \quad (3.4a)$$

$$\frac{d\nu(t)}{dt} = -\frac{1}{\sigma_X^2} \nu(t)^2, \quad (3.4b)$$

where I use $\nu(t)$ to explicitly denote that the conditional variance is a function of time. Conditional Normality follows from Theorem 11.1, and the differential equations describing the conditional moments follow from Theorem 12.1.

⁴The solution to the ODE for deterministic bonds is $B_t = e^{rt}$.

The estimated risk premium evolves stochastically with the exchange rate. To simplify the expression in (3.4), substitute the exchange rate process in (3.1) and rewrite the conditional mean as

$$\begin{aligned} d\hat{\lambda}_{Xt} &= \frac{\nu(t)}{\sigma_X} \left[\frac{\lambda_X - \hat{\lambda}_{Xt}}{\sigma_X} dt + dZ_{Xt} \right] \\ &= \frac{\nu(t)}{\sigma_X} d\hat{Z}_{Xt} \end{aligned} \quad (3.5)$$

where

$$d\hat{Z}_{Xt} = \frac{\lambda_X - \hat{\lambda}_{Xt}}{\sigma_X} dt + dZ_{Xt}, \quad (3.6)$$

is a standard Brownian motion adapted to the investor's observable information set. Using (3.6), rewrite the process for the exchange rate as

$$\frac{dX_t}{X_t} = \left[r_d - r_f + \hat{\lambda}_{Xt} \right] dt + \sigma_X d\hat{Z}_{Xt}, \quad (3.7)$$

which depends only on observable processes. In this setting, the estimated risk premium is perfectly correlated with the exchange rate. As investors observe new information on currency prices, they revise up or down their estimate of the currency risk premium.⁵

Properties of the Estimator

Theorem 12.2 from LS provides the solutions to the differential equations in (3.4), and establishes that $\hat{\lambda}_X$ is a consistent estimator of λ_X . The estimated risk premium is a stochastic function of time, and the expectation is given by

$$E \left[\hat{\lambda}_{Xt} \right] = \frac{t\nu_0\lambda_X + \hat{\lambda}_{X0}\sigma_X^2}{t\nu_0 + \sigma_X^2} \quad (3.8)$$

with

$$\lim_{t \rightarrow \infty} E \left[\hat{\lambda}_{Xt} \right] = \lambda_X.$$

In other words, the conditional mean of the posterior distribution is converging asymptotically to the true value of the risk premium.

The conditional (or posterior) variance of the estimate, $\nu(t)$, is a deterministic function of time, and satisfies an ordinary differential equation with initial condition

⁵This is not always true - in a more general setting where the foreign stock is correlated with the exchange rate, the conditional mean will be spanned by the stock and fx processes, but will not be perfectly correlated with either.

$\nu(0) = \nu_0$. The solution is

$$\nu(t) = \frac{\nu_0 \sigma_X^2}{t \nu_0 + \sigma_X^2}, \quad (3.9)$$

which declines monotonically to zero as $t \rightarrow \infty$. Thus, estimation risk disappears asymptotically as investors learn the true value of λ_X . (This is not generally true when the unobserved risk premium is stochastic.)

3.2.3 The Investor's Asset Allocation Problem

Foreign bonds and equity are denominated in foreign currency. However, a domestic investor is concerned about payoffs in domestic currency and must convert the foreign currency payoffs to domestic currency at prevailing exchange rates. Define the domestic currency prices of the foreign bond and stock as P_{Xt} and P_{St} , with $P_{Xt} = X_t B_{ft}$ and $P_{St} = X_t S_t$. An application of Ito's Lemma delivers the domestic currency dynamics of foreign assets:

$$\frac{dP_{Xt}}{P_{Xt}} = \left[r_d + \hat{\lambda}_{Xt} \right] dt + \sigma_X d\hat{Z}_{Xt} \quad (3.10a)$$

$$\frac{dP_{St}}{P_{St}} = \left[r_d + \hat{\lambda}_{Xt} + \lambda_S \right] dt + \sigma_X d\hat{Z}_{Xt} + \sigma_S dZ_{St}. \quad (3.10b)$$

While the foreign bond is risk-free in foreign currency, it is risky in domestic currency due to exchange rate risk; from the standpoint of the domestic investor, the foreign bond earns the currency risk premium in excess of the domestic risk-free rate. This excess return will introduce a speculative motive for holding foreign currency, which is otherwise absent when risk premia are zero.

When converted to domestic currency, the foreign stock is exposed to both equity risk and currency risk. Note that the foreign bond (which only carries currency risk) is perfectly correlated with the currency component of foreign equity - an investor who is long foreign equity can eliminate currency risk by taking an offsetting short position in foreign bonds. When the currency risk premium is zero and investors are myopic, a complete hedge is optimal. This is the basis for advice to completely hedge currency risk.

The investor has CRRA utility over terminal wealth, and chooses a dynamic investment strategy in the foreign stock, the foreign bond, and the domestic risk-free bond. Define α_X as the optimal allocation to foreign bonds (as a percentage of total wealth), and α_S as the optimal allocation to foreign equity. The residual allocation to domestic bonds is $1 - \alpha_X - \alpha_S$, and the ratio $-\alpha_X/\alpha_S$ determines the optimal

currency hedge. Stated formally, the optimization problem is

$$J(W, \hat{\lambda}_{Xt}, t) = \max_{\alpha_X, \alpha_S} E \left[\frac{W^{1-\gamma}}{1-\gamma} \right],$$

subject to the dynamic budget constraint,

$$\frac{dW}{W} = \left[r_d + \alpha_X \hat{\lambda}_{Xt} + \alpha_S \lambda_S \right] dt + (\alpha_X + \alpha_S) \sigma_X d\hat{Z}_{Xt} + \alpha_S \sigma_S dZ_{St}.$$

To solve the investor's portfolio problem, I use the dynamic programming approach. Formally, this is a dynamic portfolio problem with a stochastic investment opportunity set, with variation in the opportunity set summarized by the stochastic properties of the estimated risk premium. (In other words, the estimated risk premium is a state variable.) This is true even though the unobserved investment opportunity set is constant; the investor is learning about the risk premium while simultaneously making a portfolio decision, and uses the estimated risk premium (which is only non-stochastic asymptotically) to solve the asset allocation problem.

The investor's indirect utility function is $J(W, \hat{\lambda}_{Xt}, t)$; the Bellman principle of optimality implies $E[dJ] = 0$. Applying Ito's Lemma to J and taking expectations yields the Hamilton-Jacobi-Bellman equation for this problem:

$$0 = \max_{\alpha_X, \alpha_S} \left\{ J_W W \left[r_d + \alpha_X \hat{\lambda}_{Xt} + \alpha_S (\hat{\lambda}_{Xt} + \lambda_S) \right] + \frac{1}{2} J_{WW} W^2 \left[\alpha_S^2 \sigma_S^2 + (\alpha_S + \alpha_X)^2 \sigma_X^2 \right] + \frac{1}{2} J_{\hat{\lambda}\hat{\lambda}} \frac{\nu(t)^2}{\sigma_X^2} + J_{W\hat{\lambda}} W \nu(t) (\alpha_X + \alpha_S) + \frac{\partial J}{\partial t} \right\} \quad (3.11)$$

The optimal risky asset demands follow from the first order conditions:

$$\alpha_X = -\frac{J_W}{J_{WW} W} \left(\frac{\hat{\lambda}_{Xt}}{\sigma_X^2} - \frac{\lambda_S}{\sigma_S^2} \right) - \frac{J_{W\hat{\lambda}}}{J_{WW} W} \frac{\nu(t)}{\sigma_X^2} \quad (3.12a)$$

$$\alpha_S = -\frac{J_W}{J_{WW} W} \frac{\lambda_S}{\sigma_S^2} \quad (3.12b)$$

The optimal allocation to foreign equity is determined by the risk-return tradeoff of the stock, as well as the coefficient of relative risk aversion. The allocation is myopic because the state variable is uncorrelated with stock returns; if the correlation were non-zero, there would be an intertemporal hedging term as well.

The optimal allocation to foreign bonds is made up of a myopic allocation and an intertemporal hedging term. The myopic allocation contains two components: The speculative demand for foreign currency, which is non-zero as long as the estimated risk premium is non-zero, and the foreign currency hedge. The second term enters the

myopic portfolio due the currency risk of the foreign stock. The optimal allocation to foreign equity is determined by the equity risk-return tradeoff - however, when purchasing foreign stock, a domestic investor takes on a potentially sub-optimal level of currency risk. The investor offsets, or hedges, this risk with a short position in foreign bonds, which are exposed only to currency risk.⁶

The intertemporal hedging term in the demand for foreign bonds is due to parameter uncertainty and learning. It is zero when there is no estimation risk ($\nu(t) = 0$), or when there is no motive to hedge ($J_{W\hat{\lambda}} = 0$, as is the case with logarithmic utility).

The optimal currency hedge is defined as the negative of the ratio of foreign bond demand to foreign equity demand,

$$-\frac{\alpha_X}{\alpha_S} = 1 - \frac{\hat{\lambda}_{Xt}\sigma_S^2}{\lambda_S\sigma_X^2} - \frac{J_{W\hat{\lambda}}\nu(t)\sigma_S^2}{J_W\lambda_S\sigma_X^2}. \quad (3.13)$$

When there is no parameter uncertainty, the optimal hedge is independent of the investor's level of risk aversion. This is because, without parameter uncertainty, investors are myopic and invest in the tangency portfolio - while risk aversion will determine the optimal investment in the tangency portfolio vs. risk-free bonds, the weights within the tangency portfolio are fixed and independent of risk aversion.

In addition to full information, if the foreign currency risk premium is zero, the optimal hedge is a complete one: investors offset 100% of the currency risk of foreign equity with a short position in foreign bonds. This is the result of Solnik (1974). In the presence of estimation risk, the optimal currency hedge will depend on, among other things, risk aversion.

3.2.4 Value Function Solution

Evaluated at the optimum, the Bellman equation in (3.11) describes a partial differential equation for expected utility. To solve the PDE, I follow the approach of Kim and Omberg (1996) and conjecture the following solution:⁷

$$J(W, \hat{\lambda}_{Xt}, t) = \frac{W^{1-\gamma}}{1-\gamma} \exp \left\{ (1-\gamma) \left(A(t) + \frac{1}{2}C(t)\hat{\lambda}_{Xt}^2 \right) \right\} \quad (3.14)$$

where $A(t)$ and $C(t)$ are deterministic functions of time. Substitution of the conjecture and its derivatives into (3.12) and (3.11) results in an equation that is quadratic in $\hat{\lambda}_{Xt}$, the conditional mean; its coefficients must be zero for the PDE to hold. Set-

⁶If the foreign equity risk premium is negative, so that the optimal equity allocation is a short position, then the offsetting currency hedge will be a long position in foreign bonds.

⁷The general solution for this problem is $J(W, \hat{\lambda}_{Xt}, t) = \frac{W^{1-\gamma}}{1-\gamma} \exp \left\{ (1-\gamma) \left(A(t) + B(t)\hat{\lambda}_{Xt} + \frac{1}{2}C(t)\hat{\lambda}_{Xt}^2 \right) \right\}$, with $B[T] = 0$. However, when the exchange rate and foreign stock are uncorrelated, it is straightforward to verify that $B(t) = 0$ for all t .

ting the coefficients equal to zero results in a recursive system of ordinary differential equations:

$$\frac{dC(t)}{dt} = -\frac{1}{\gamma\sigma_X^2} - \frac{2(1-\gamma)\nu(t)}{\gamma\sigma_X^2}C(t) - \frac{1-\gamma}{\gamma}\frac{\nu(t)^2}{\sigma_X^2}C(t)^2$$

$$\frac{dA(t)}{A(t)} = -r_d - \frac{1}{2}\frac{\lambda_S^2}{\sigma_S^2} - \frac{1}{2}\frac{\nu(t)^2}{\sigma_X^2}C(t)$$

with boundary conditions $A(T) = C(T) = 0$ and $\nu(t)$ as defined in (3.9). The solutions to the ODEs are given by:

$$C(t) = \frac{(T-t)(t\nu_0 + \sigma_X^2)}{\sigma_X^2((t+T(\gamma-1))\nu_0 + \gamma\sigma_X^2)}$$

$$A(t) = \frac{(T-t)(\lambda_S^2 + 2\gamma r_d \sigma_S^2)}{2\gamma\sigma_S^2} + \frac{\left(\ln\left[\frac{T\nu_0 + \sigma_X^2}{t\nu_0 + \sigma_X^2}\right] + \gamma \ln\left[1 - \frac{(T-t)\nu_0}{\gamma(T\nu_0 + \sigma_X^2)}\right]\right)}{2(\gamma-1)}$$

Finally, substitution of the value function solution into (3.13) yields the optimal hedge ratio:

$$-\frac{\alpha_X}{\alpha_S} = 1 - \frac{\hat{\lambda}_{Xt}\sigma_S^2}{\lambda_S\sigma_X^2} \left[1 - \left(1 - \frac{1}{\gamma}\right) \frac{(T-t)\nu_0}{\left(t+T\left(1 - \frac{1}{\gamma}\right)\right)\nu_0 + \sigma_X^2} \right]. \quad (3.17)$$

3.2.5 Analysis

When $t = 0$, the currency hedge ratio simplifies to

$$-\frac{\alpha_X}{\alpha_S} = 1 - \frac{\hat{\lambda}_{X0}\sigma_S^2}{\lambda_S\left(T\left(1 - \frac{1}{\gamma}\right)\nu_0 + \sigma_X^2\right)}. \quad (3.18)$$

Figure 3.1 plots the optimal hedge rule as a function the investment horizon, for various levels of risk aversion.⁸ The estimated currency risk premium is $\hat{\lambda}_{X0} = .005$, and the prior variance is also $\nu_0 = .005$ (or a standard deviation of approximately .071). The remaining values are .05 for the equity risk premium, and .10 and .15 for the exchange rate and equity volatilities, respectively.

Inspection of both Figure 3.1 and equation (3.18) reveals that, for any level of risk

⁸I choose $\gamma = 1.1$ for the lowest level to illustrate an investor who is close to a logarithmic investor. When utility is logarithmic ($\gamma = 1$), the optimal hedge is independent of the investment horizon.

aversion greater than logarithmic (i.e. $\gamma > 1$), increasing variance and investment horizon will eventually outweigh the speculative motive for holding foreign currency, and the optimal hedge approaches a complete one.

However, this is not the case for risk aversion. Taking limits,

$$\lim_{\gamma \rightarrow \infty} -\frac{\alpha_X}{\alpha_S} = 1 - \frac{\hat{\lambda}_{Xt}\sigma_S}{\lambda_S(T\nu_0 + \sigma_X^2)},$$

which will not equal 1 unless the conditional risk premium is zero.⁹ To understand this result, it is useful to rewrite the optimal demand for foreign bonds:

$$\alpha_X = \frac{1}{\gamma} \left(\frac{\hat{\lambda}_{Xt}}{T \left(1 - \frac{1}{\gamma}\right) \nu_0 + \sigma_X^2} \right) - \frac{1}{\gamma} \frac{\lambda_S}{\sigma_S^2} \quad (3.19)$$

The first term on the right hand side can be viewed as the adjusted speculative demand for foreign currency, with the risk aversion parameter controlling how much an investor adjusts for estimation risk. Speculative demand for foreign currency is determined by the risk-return tradeoff, and risk aversion determines how an investor views this tradeoff. Intuitively, uncertainty about the risk premium, and therefore uncertainty about currency returns, gets added to the variance of returns. A myopic investor will optimally ignore parameter uncertainty, while an infinitely risk averse investor will fully include parameter uncertainty in the foreign exchange variance.

3.3 A General Model of Portfolio Choice

In this section I present a portfolio choice model that generalizes the simple model presented earlier. The basic framework is the same, except that I allow for non-zero correlation between the foreign exchange rate and foreign stock, and also introduce a stochastic risk premium. For simplicity, interest rates remain constant. The simple model is a special case of the general model presented here.

3.3.1 Asset Market Dynamics

The processes for the foreign exchange rate and foreign equity are identical to (3.1) and (3.2), except that the processes are correlated, with the correlation equal to ρ_{XS} . For the general model, it is convenient to stack the two processes into a vector \mathcal{I}_t , which I call the investor's information set:

⁹Although the optimal hedge limits towards a constant value (holding time and estimation risk fixed), an infinitely risk averse investor will invest his entire wealth in domestic bonds, so no currency hedging actually takes place.

$$\begin{aligned}
d\mathcal{I}_t &= \begin{bmatrix} \frac{dX_t}{X_t} \\ \frac{dS_t}{S_t} \end{bmatrix} \\
&= \left[\begin{pmatrix} r_d - r_f \\ r_f + \lambda_S \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \lambda_{Xt} \right] dt + \begin{bmatrix} \sigma_X & 0 \\ \sigma_S \rho_{XS} & \sigma_S \sqrt{1 - \rho_{XS}^2} \end{bmatrix} \times \begin{bmatrix} dZ_{Xt} \\ dZ_{St} \end{bmatrix} \\
&= [\mathbf{M}_0(\mathcal{I}, t) + \mathbf{M}_1(\mathcal{I}, t)\lambda_{Xt}] dt + \Sigma(\mathcal{I}, t)d\mathbf{Z}_{\mathcal{I}}.
\end{aligned} \tag{3.20}$$

where λ_{Xt} is the unobservable (stochastic) risk premium, and $\mathbf{Z}_{\mathcal{I}}$ is an unobservable, two-dimensional Brownian motion. As with the simple model, investors will use their observations of the exchange rate and stock returns to form an optimal estimate of the risk premium. Because the two processes are correlated, equity returns will contain information about the risk premium which investors can use to improve their estimate.

The risk premium is stochastic and evolves according to an Ornstein-Uhlenbeck process:

$$d\lambda_{Xt} = \kappa_\lambda \left(\frac{1}{2} \sigma_X^2 - \lambda_{Xt} \right) dt + \Sigma'_\lambda d\mathbf{Z}_{\mathcal{I}} + \omega_\lambda dZ_\lambda, \tag{3.21}$$

with

$$\begin{aligned}
\Sigma_\lambda &= [\sigma_\lambda \rho_{X\lambda} \quad \sigma_\lambda \hat{\rho}_{S\lambda}]' \\
\omega_\lambda &= \sigma_\lambda \sqrt{1 - \rho_{X\lambda}^2 - \hat{\rho}_{S\lambda}^2} \\
\hat{\rho}_{S\lambda} &= \frac{\rho_{S\lambda} - \rho_{XS} \rho_{X\lambda}}{\sqrt{1 - \rho_{XS}^2}},
\end{aligned}$$

where Z_λ is a standard Brownian motion orthogonal to Z_X and Z_S , $\rho_{X\lambda}$ is the correlation between foreign exchange and the risk premium, and $\rho_{S\lambda}$ is the correlation between foreign equity and the risk premium. In general, the risk premium will not be spanned by foreign exchange and equity, and markets are incomplete. When $\omega_\lambda = 0$, markets are complete.¹⁰ When $\kappa_\lambda = \sigma_\lambda = 0$, the risk premium is constant as in Section 3.2.

The risk premium reverts towards a long-run mean equal to $\frac{1}{2}\sigma_X^2$, with the rate of mean reversion governed by the parameter κ_λ . To understand this particular choice for the long-run mean, it is useful to briefly consider the problem of a foreign investor. While a domestic investor prices foreign assets in terms of domestic currency, and is

¹⁰The assumption of complete markets imposes additional restrictions on the correlations.

therefore interested in the exchange rate X_t , a foreign investor will be interested in the foreign price of domestic currency, or $Y_t = 1/X_t$. An application of Ito's Lemma delivers the dynamics of Y_t ,

$$\frac{dY_t}{Y_t} = [r_f - r_d + \lambda_{Yt}] dt - \sigma_X dZ_{Xt} \quad (3.23a)$$

$$\lambda_{Yt} = \sigma_X^2 - \lambda_{Xt}, \quad (3.23b)$$

with

$$d\lambda_{Yt} = \kappa_\lambda \left(\frac{1}{2} \sigma_X^2 - \lambda_{Yt} \right) dt - \Sigma'_\lambda d\mathbf{Z}_\mathcal{I} - \omega_\lambda dZ_\lambda.$$

Notice from (3.23) that exchange rate risk premia do not sum to zero, as intuition would suggest, but rather to the variance of the exchange rate. This is true for means as well, and risk premia are identical when $\lambda_{Xt} = \lambda_{Yt} = \frac{1}{2} \sigma_X^2$. This makes $\frac{1}{2} \sigma_X^2$ a convenient choice for the long-run mean since it yields symmetric exchange rate processes for foreign and domestic investors.

Economically, when the foreign exchange risk premium is equal to its long-run mean, foreign and domestic investors both expect a positive excess return from speculation in foreign currency. This is a manifestation of Siegel's Paradox (Siegel; 1972) which, roughly speaking, states that the exchange rate risk premium cannot be simultaneously zero for both foreign and domestic investors, and is due to Jensen's Inequality.

In other words, the textbook case of uncovered interest parity (UIP), which assumes excess speculative returns are zero, is not consistent with symmetric exchange rate processes; it can only be the case that risk premia are jointly zero if foreign exchange rates are deterministic. However, log exchange rates can simultaneously satisfy a zero-risk-premium condition, which can be seen by applying Ito's Lemma to $\log X_t$ and $\log Y_t$:

$$\begin{aligned} d\log X_t &= \left[r_d - r_f + \lambda_{Xt} - \frac{1}{2} \sigma_X^2 \right] dt + \sigma_X dZ_{Xt} \\ d\log Y_t &= \left[r_f - r_d + \lambda_{Yt} - \frac{1}{2} \sigma_X^2 \right] dt - \sigma_X dZ_{Xt} \end{aligned}$$

When $\lambda_{Xt} = \lambda_{Yt} = \frac{1}{2} \sigma_X^2$, UIP holds for log exchange rates. For this reason, I refer to the long-run mean case, with positive currency risk premia for both foreign and domestic investors, as the UIP case.

3.3.2 The Inference Problem

Theorem 12.7 from Liptser and Shirayev (1978) provides the multidimensional counterpart to Theorem 12.1. The investor's optimal estimate of the risk premium

is $\hat{\lambda}_{Xt}$ with variance ν_t , and conditional Normality of the posterior distribution again applies (assuming a Normal prior). The differential equations describing the laws of motion for the first and second moments are:

$$\begin{aligned} d\hat{\lambda}_{Xt} = & \kappa_\lambda \left(\frac{1}{2}\sigma_X^2 - \hat{\lambda}_{Xt} \right) dt + [\Sigma'_\lambda \Sigma'_\mathcal{I} + \nu(t)\mathbf{M}'_1] \\ & \times (\Sigma_\mathcal{I} \Sigma'_\mathcal{I})^{-1} \left[d\mathcal{I}_t - (\mathbf{M}_0 + \mathbf{M}_1 \hat{\lambda}_{Xt}) dt \right] \end{aligned} \quad (3.25a)$$

$$\begin{aligned} \frac{d\nu(t)}{dt} = & -2\kappa_\lambda \nu(t) + \Sigma'_\lambda \Sigma_\lambda + \omega_\lambda^2 - [\Sigma'_\lambda \Sigma'_\mathcal{I} + \nu(t)\mathbf{M}'_1] \\ & \times (\Sigma_\mathcal{I} \Sigma'_\mathcal{I})^{-1} [\Sigma'_\lambda \Sigma'_\mathcal{I} + \nu(t)\mathbf{M}'_1]' \end{aligned} \quad (3.25b)$$

To simplify the expression for the conditional mean, define a new Brownian motion $\hat{\mathbf{Z}}_\mathcal{I}$, which is observed by investors:

$$\begin{aligned} d\hat{\mathbf{Z}}_\mathcal{I} = & \Sigma_\mathcal{I}^{-1} \left[d\mathcal{I}_t - (\mathbf{M}_0 + \mathbf{M}_1 \hat{\lambda}_{Xt}) dt \right] \\ = & \Sigma_\mathcal{I}^{-1} \mathbf{M}_1 \left(\lambda_{Xt} - \hat{\lambda}_{Xt} \right) dt + d\mathbf{Z}_\mathcal{I}. \end{aligned} \quad (3.26)$$

Using this expression, rewrite the process for estimated risk premium as,

$$d\hat{\lambda}_{Xt} = \kappa_\lambda \left(\frac{1}{2}\sigma_X^2 - \hat{\lambda}_{Xt} \right) dt + [\Sigma_\lambda + \nu(t)\Sigma_\mathcal{I}^{-1} \mathbf{M}_1]' d\hat{\mathbf{Z}}_\mathcal{I}, \quad (3.27)$$

and the information set as,

$$d\mathcal{I}_t = \left[\mathbf{M}_0 + \mathbf{M}_1 \hat{\lambda}_{Xt} \right] dt + \Sigma_\mathcal{I} d\hat{\mathbf{Z}}_\mathcal{I}, \quad (3.28)$$

which depend only on observable parameters. Furthermore, although markets are incomplete, markets appear dynamically complete to investors since the estimated risk premium is spanned by the information set.

As in the simple model, the conditional (or posterior) variance of the estimate, $\nu(t)$, is a deterministic function of time, and satisfies a Ricatti ordinary differential equation; the expression in (3.25) simplifies to:

$$\begin{aligned} \frac{d\nu(t)}{dt} = & \omega_\lambda^2 - 2 \left(\kappa_\lambda + \Sigma'_\lambda \Sigma_\mathcal{I}^{-1} \mathbf{M}_1 \right) \nu(t) - \mathbf{M}'_1 (\Sigma_\mathcal{I} \Sigma'_\mathcal{I})^{-1} \mathbf{M}_1 \nu(t)^2 \\ = & \omega_\lambda^2 - 2 \left(\kappa_\lambda + \frac{\rho_{X\lambda} - \rho_{XS} \rho_{S\lambda}}{1 - \rho_{XS}^2} \frac{\sigma_\lambda}{\sigma_X} \right) \nu(t) - \left(\frac{1}{1 - \rho_{XS}^2} \frac{1}{\sigma_X^2} \right) \nu(t)^2. \end{aligned} \quad (3.29)$$

The ODE in (3.29) has a closed-form solution, discussed in the Appendix B. The

variance declines monotonically in time, to a long-run (asymptotic) value. Heuristically, the long-run value is found by setting the right-hand side of (3.29) equal to zero. This steady-state value will in general be non-zero, implying that investors never fully learn the true stochastic process of the risk premium. Therefore, the estimator is generally inconsistent. The intuition behind this result is that investors are trying to learn about a process which is itself stochastic and not perfectly spanned by observable processes. As a result, investors are never able to perfectly filter the risk premium, since each time they form an estimate, the true risk premium is moving in a way they cannot observe.¹¹

Figure 3.2 plots a sample solution to the variance ODE when all correlations are zero, the prior variance is $\nu_0 = .005$, and the volatilities of the risk premium and exchange rate are $\sigma_\lambda = .05$ and $\sigma_X = .10$.

3.3.3 The Portfolio Problem

As in the simple model, a risk-averse investor chooses an optimal portfolio of foreign stocks and bonds to maximize expected utility over terminal wealth. The domestic currency prices of foreign assets follow from Ito's Lemma. As with the information set, it is convenient to stack the prices into a vector, which I call $d\mathbf{P}_t$:

$$d\mathbf{P}_t = \begin{bmatrix} \frac{dP_{Xt}}{P_{Xt}} \\ \frac{dP_{St}}{P_{St}} \end{bmatrix} = [r_d \mathbf{1} + \mathbf{\Lambda}_P] dt + \mathbf{\Sigma}_P d\hat{\mathbf{Z}}_T, \quad (3.30)$$

with

$$\mathbf{\Lambda}_P = \begin{bmatrix} \hat{\lambda}_{Xt} \\ \hat{\lambda}_{Xt} + \lambda_S + \sigma_X \sigma_S \rho_{XS} \end{bmatrix}, \quad \mathbf{\Sigma}_P = \begin{bmatrix} \sigma_X & 0 \\ \sigma_X + \sigma_S \rho_{XS} & \sigma_S \sqrt{1 - \rho_{XS}^2} \end{bmatrix}.$$

The optimal allocation to foreign assets is the 2-by-1 vector $\boldsymbol{\alpha}$; the first element (α_X) is the allocation to foreign bonds, and the second element (α_S) is the optimal allocation to foreign equity. The optimal currency hedge is again given by $-\alpha_X/\alpha_S$, and the allocation to domestic bonds is $1 - \boldsymbol{\alpha}'\mathbf{1}$.

The optimization problem is

$$J(W, \hat{\lambda}_{Xt}, t) = \max_{\boldsymbol{\alpha}} E \left[\frac{W^{1-\gamma}}{1-\gamma} \right],$$

subject to the dynamic budget constraint,

$$\frac{dW}{W} = [r_d + \boldsymbol{\alpha}'\mathbf{\Lambda}_P] dt + \boldsymbol{\alpha}'\mathbf{\Sigma}_P d\hat{\mathbf{Z}}_T.$$

¹¹Even when the stochastic process is perfectly spanned by the information set, the estimator may be inconsistent. See Gennotte (1986) for a discussion.

The Bellman equation for this problem is

$$\begin{aligned}
0 = \max_{\boldsymbol{\alpha}} \left\{ \frac{\partial J}{\partial t} + J_W W [r_d + \boldsymbol{\alpha}' \boldsymbol{\Lambda}_P] + \frac{1}{2} J_{WW} W^2 \boldsymbol{\alpha}' \boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P \boldsymbol{\alpha} + \kappa_\lambda \left(\frac{1}{2} \sigma_X^2 - \hat{\lambda}_{Xt} \right) J_{\hat{\lambda}} \right. \\
+ \frac{1}{2} (\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \mathbf{M}_1)' (\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \mathbf{M}_1) J_{\hat{\lambda}\hat{\lambda}} \\
\left. + W \boldsymbol{\alpha}' \boldsymbol{\Sigma}_P (\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \mathbf{M}_1) J_{W\hat{\lambda}} \right\}.
\end{aligned} \tag{3.31}$$

The investor's optimal asset allocation strategy follows from the first-order condition with respect to $\boldsymbol{\alpha}$,

$$\begin{aligned}
\boldsymbol{\alpha} = \begin{bmatrix} \alpha_X \\ \alpha_S \end{bmatrix} = - \frac{J_W}{J_{WW} W} (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P)^{-1} \boldsymbol{\Lambda}_P - \frac{J_{W\hat{\lambda}}}{J_{WW} W} (\boldsymbol{\Sigma}'_P)^{-1} \boldsymbol{\Sigma}_\lambda \\
- \frac{J_{W\hat{\lambda}}}{J_{WW} W} (\boldsymbol{\Sigma}'_P)^{-1} \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \mathbf{M}_1 \nu(t).
\end{aligned} \tag{3.32}$$

The first term on the right hand side is the investor's myopic portfolio demand, which is determined by the mean-variance tradeoff for foreign stocks and bonds. The second and third terms describe the intertemporal hedging demand due to the stochastic risk premium. This portfolio has two components: The first component hedges against variation in the investment opportunity set due to the stochastic properties of the unobserved risk premium; it is zero when the unobserved risk premium is constant (no need to hedge), when all the correlations are zero (no ability to hedge), and when $J_{W\hat{\lambda}}$ is zero (no motive to hedge). The second component is due to parameter uncertainty and learning; it is zero when $J_{W\hat{\lambda}}$ is zero, when there is no parameter uncertainty ($\nu(t) = 0$), and when there is no learning.

3.3.4 Value Function

The value function for the general problem takes the same form as in the simple problem,

$$J(W, \hat{\lambda}_{Xt}, t) = \frac{W^{1-\gamma}}{1-\gamma} e^{(1-\gamma)(A(t)+B(t)\hat{\lambda}_{Xt}+\frac{1}{2}C(t)\hat{\lambda}_{Xt}^2)}$$

where $A(t)$, $B(t)$, and $C(t)$ satisfy a system of ordinary differential equations given in Appendix B. Using this solution, the optimal portfolio rule is

$$\boldsymbol{\alpha} = \frac{1}{\gamma} (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P)^{-1} \boldsymbol{\Lambda}_P - \left(1 - \frac{1}{\gamma} \right) (\boldsymbol{\Sigma}'_P)^{-1} (\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \mathbf{M}_1) \left(B(t) + C(t) \hat{\lambda}_{Xt} \right) \tag{3.33}$$

where the first term is the myopic portfolio demand and the second term is the intertemporal hedging demand.

The optimal currency hedge is again given by the ratio of foreign bonds to foreign stock. When all correlations are equal to zero, the hedge ratio reduces to,

$$-\frac{\alpha_X}{\alpha_S} = 1 - \frac{\hat{\lambda}_{Xt}\sigma_S^2}{\lambda_S\sigma_X^2} \left[1 - (\gamma - 1)C(t)\nu(t) \right] - (1 - \gamma)\frac{\sigma_S^2}{\lambda_S\sigma_X^2}B(t)\nu(t). \quad (3.34)$$

It is generally not possible to reduce the hedge ratio further, since the functions $B(t)$ and $C(t)$ will be available in closed form only in certain cases. However, they are easily solved numerically.

3.4 Numerical Analysis

In the absence of parameter uncertainty, a time-varying risk premium leads to substantial market timing in the optimal hedge. When the risk premium is positive, investors will reduce their hedge ratio below 1 in response to the positive expected return to currency speculation. Likewise, when the risk premium is negative, the optimal speculative policy will be to short foreign currency, so the hedge ratio will be greater than 1. This is also true for myopic investors.

However, as Figure 3.3 shows, estimation risk considerably dampens the speculative component for an investor with a five year investment horizon. The optimal currency hedge much closer to a complete one, especially for values of the risk premium close to zero. The parameter values are identical to the simple model; in addition, the rate of mean reversion is .25, the volatility of the unobserved risk premium is .05, and all correlations are set to zero.¹²

As with the simple model presented earlier, there is a strong horizon effect at work. Figure 3.4 plots the optimal hedge ratio as a function of investment horizon for four values of risk aversion. I use $\gamma = 1.1$ as the first value to again show the policy of an investor who is close to myopic.

As with the simple case, the hedging policy is converging towards a long-run value as the horizon increases, and the value is closer to 1 as risk aversion increases. Notice however that the long-run hedge does not in general converge towards a complete hedge; this is due to the third term in (3.34), which does not depend on the value of the estimated risk premium. Whereas in the simple model, increasing time horizon and prior variance lead an investor to act as if their estimate of the risk premium is zero (in which case the optimal hedge is a complete one), this is not true in the

¹²As is apparent from (3.34), the optimal hedging rule is linear in the risk premium when all stochastic processes are uncorrelated. When processes are correlated, the hedging rule will still be a monotonic function of the risk premium; however, the function may either be convex or concave depending on the signs of the correlations.

general model.

3.5 The Cost of Suboptimal Hedging

Given that, in a world of parameter uncertain, optimal hedge ratios approach a complete hedge as the investment horizon increases, a natural question to ask deals with the economic importance of estimation risk. In particular, what is the opportunity cost of following a suboptimal strategy, in terms of certainty equivalence? How much extra wealth would an investor require to follow a suboptimal strategy such as a complete hedge?

3.5.1 Expected Utility of Terminal Wealth

In this section I return to the simple model of portfolio choice in Section 3.2, but I make use of the vector notation introduced in Section 3.3. The analysis closely follows that of Larsen and Munk (2010). To begin, consider any portfolio strategy $\hat{\alpha}$. The dynamic budget constraint is

$$\frac{dW}{W} = [r_d + \hat{\alpha}' \Lambda_P] dt + \hat{\alpha}' \Sigma_P d\hat{Z}_{\mathcal{I}}$$

so that total wealth at time T is given by

$$W_T = W_t \exp \left\{ \int_t^T \left[r_d + \hat{\alpha}' \Lambda_P - \frac{1}{2} \hat{\alpha}' \Sigma_P \Sigma_P' \hat{\alpha} \right] ds + \int_t^T \hat{\alpha}' \Sigma_P d\hat{Z}_{\mathcal{I}} \right\}. \quad (3.35)$$

Then expected utility of terminal wealth is given by

$$J(W, \hat{\lambda}_{Xt}, t) = E_t \left[\frac{W_t^{1-\gamma}}{1-\gamma} \right] H(\hat{\lambda}_{Xt}, t) \quad (3.36)$$

with

$$H(\hat{\lambda}_{Xt}, t) = E_t \left[\exp \left\{ (1-\gamma) \int_t^T \left[r_d + \hat{\alpha}' \Lambda_P - \frac{1}{2} \hat{\alpha}' \Sigma_P \Sigma_P' \hat{\alpha} \right] ds + (1-\gamma) \int_t^T \hat{\alpha}' \Sigma_P d\hat{Z}_{\mathcal{I}} \right\} \right]. \quad (3.37)$$

Next, define η_τ such that

$$\eta_\tau = \exp \left\{ (1-\gamma) \int_t^\tau \hat{\alpha}' \Sigma_P d\hat{Z}_{\mathcal{I}} - \frac{1}{2} (1-\gamma)^2 \int_t^\tau \hat{\alpha}' \Sigma_P \Sigma_P' \hat{\alpha} ds \right\}. \quad (3.38)$$

Using Girsanov's Theorem, this allows us to rewrite (3.37) as

$$\begin{aligned} H(\hat{\lambda}_{Xt}, t) &= E_t \left[\eta_T \exp \left\{ (1 - \gamma) \int_t^T \left[r_d + \hat{\boldsymbol{\alpha}}' \boldsymbol{\Lambda}_P - \frac{\gamma}{2} \hat{\boldsymbol{\alpha}}' \boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P \hat{\boldsymbol{\alpha}} \right] ds \right\} \right] \\ &= E_t^{\mathbb{Q}} \left[\exp \left\{ (1 - \gamma) \int_t^T \left[r_d + \hat{\boldsymbol{\alpha}}' \boldsymbol{\Lambda}_P - \frac{\gamma}{2} \hat{\boldsymbol{\alpha}}' \boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P \hat{\boldsymbol{\alpha}} \right] ds \right\} \right], \end{aligned} \quad (3.39)$$

where \mathbb{Q} is an equivalent probability measure and $d\hat{\boldsymbol{Z}}_{\mathcal{I}}^{\mathbb{Q}} = d\hat{\boldsymbol{Z}}_{\mathcal{I}} - (1 - \gamma) \boldsymbol{\Sigma}'_P \hat{\boldsymbol{\alpha}} dt$ is a standard Brownian motion under \mathbb{Q} .

Under the measure \mathbb{Q} , the estimated risk premium has dynamics,

$$d\hat{\lambda}_{Xt} = \nu(t)(1 - \gamma) \hat{\boldsymbol{\alpha}}' \boldsymbol{\Sigma}_P \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \boldsymbol{M}_1 dt + \nu(t) (\boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \boldsymbol{M}_1)' d\hat{\boldsymbol{Z}}_{\mathcal{I}}^{\mathbb{Q}}. \quad (3.40)$$

Therefore, by use of the Feynman-Kac Theorem, the function H must satisfy the partial differential equation

$$\begin{aligned} 0 &= \frac{\partial H}{\partial t} + \left[(1 - \gamma) \left(r_d + \hat{\boldsymbol{\alpha}}' \boldsymbol{\Lambda}_P - \frac{\gamma}{2} \hat{\boldsymbol{\alpha}}' \boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P \hat{\boldsymbol{\alpha}} \right) \right] H(\hat{\lambda}_{Xt}, t) \\ &\quad + \left[\nu(t)(1 - \gamma) \hat{\boldsymbol{\alpha}}' \boldsymbol{\Sigma}_P \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \boldsymbol{M}_1 \right] H_{\lambda} + \frac{1}{2} \nu(t)^2 \boldsymbol{M}_1' (\boldsymbol{\Sigma}_{\mathcal{I}} \boldsymbol{\Sigma}'_{\mathcal{I}})^{-1} \boldsymbol{M}_1 H_{\lambda\lambda}. \end{aligned} \quad (3.41)$$

with $H(\hat{\lambda}_{Xt}, T) = 1$. Using the scalar notation of Section 3.2, we can rewrite the differential equation as

$$\begin{aligned} 0 &= \frac{\partial H}{\partial t} + [(1 - \gamma)\nu(t)(\alpha_X + \alpha_s)] H_{\lambda} + \frac{1}{2} \frac{\nu(t)^2}{\sigma_X^2} H_{\lambda\lambda} \\ &\quad \left[(1 - \gamma) \left(r_d + \alpha_X \hat{\lambda}_{Xt} + \alpha_s (\hat{\lambda}_{Xt} + \lambda_s) \right. \right. \\ &\quad \left. \left. - \frac{\gamma}{2} (\alpha_s^2 \sigma_s^2 + (\alpha_s + \alpha_X)^2 \sigma_X^2) \right) \right] H(\hat{\lambda}_{Xt}, t) \end{aligned} \quad (3.42)$$

This differential equation must hold for any portfolio choice α_X and α_s . When portfolios are chosen optimally, the solution is given in equation (3.14),

$$H(\hat{\lambda}_{Xt}, t) = \exp \left\{ (1 - \gamma) \left(A(t) + \frac{1}{2} C(t) \hat{\lambda}_{Xt}^2 \right) \right\} \quad (3.43)$$

3.5.2 The Cost of Complete Hedging

Suppose that, instead of following the optimal portfolio rule, an investor chooses to completely hedge her foreign currency exposure. That is, she chooses portfolios

$$\alpha_X = -\frac{1}{\gamma} \frac{\lambda_S}{\sigma_S^2} \quad (3.44a)$$

$$\alpha_S = \frac{1}{\gamma} \frac{\lambda_S}{\sigma_S^2}. \quad (3.44b)$$

In this case, the differential equation in (3.42) simplifies to

$$0 = \frac{\partial H}{\partial t} + \left[(1 - \gamma) \left(r_d + \frac{1}{2\gamma} \frac{\lambda_S^2}{\sigma_S^2} \right) \right] H(\hat{\lambda}_{Xt}, t) + \frac{1}{2} \frac{\nu(t)^2}{\sigma_X^2} H_{\lambda\lambda}. \quad (3.45)$$

It is straightforward to verify that the solution in this case is given by

$$\begin{aligned} H(\hat{\lambda}_{Xt}, t) &= \exp \{ (1 - \gamma) B(t) \} \\ B(t) &= (T - t) \left[r_d + \frac{1}{2\gamma} \frac{\lambda_S^2}{\sigma_S^2} \right] \end{aligned} \quad (3.46)$$

so that expected utility over terminal wealth is

$$J(W, \hat{\lambda}_{Xt}, t) = \frac{W_t^{1-\gamma}}{1-\gamma} \exp \{ (1 - \gamma) B(t) \} \quad (3.47)$$

As shown in Section 3.2, a complete currency hedge is suboptimal, though less so for investors with long horizons. Define L as the percentage of initial wealth an investor must receive in order to be indifferent between following the optimal strategy with initial wealth W_0 , or a suboptimal complete hedging strategy with initial wealth $(1 + L)W_0$. The investor is indifferent between the two strategies if they yield the same expected utility, i.e.

$$\frac{[(1 + L)W_0]^{1-\gamma}}{1-\gamma} \exp \{ (1 - \gamma) B(t) \} = \frac{W_0^{1-\gamma}}{1-\gamma} \exp \left\{ (1 - \gamma) \left(A(t) + \frac{1}{2} C(t) \hat{\lambda}_{Xt}^2 \right) \right\}. \quad (3.48)$$

Since L is the additional wealth necessary to set the expected utility from following a suboptimal strategy equal to the expected utility from following the optimal strategy, we can think of L as the utility loss, or opportunity cost, of following a suboptimal strategy. Solving for L , the opportunity cost of following a complete hedging strategy is given by

$$L = \exp \left\{ A(t) - B(t) + \frac{1}{2} C(t) \hat{\lambda}_{Xt}^2 \right\} \quad (3.49)$$

Figure 3.5 plots the opportunity cost of complete hedging for up to a horizon of 10 years, for three levels of risk aversion. The top figure gives the loss as a percentage of initial wealth, while the bottom figure converts the loss into an annuity stream. Intuitively, we can think of the annuity stream as an annualized opportunity cost, or the annual increase in wealth an investor would need to receive to make her indifferent between complete hedging and optimal hedging.¹³

The figures reveal two insights. First, even though optimal currency hedges approach a complete hedge as the investment horizon increases, utility losses increase with the horizon. The reason for this is that, by engaging in a complete hedge, the investor is ignoring the benefits of learning. As time evolves and the investor learns about the true value of the risk premium, her optimal hedge will gradually move away from a complete hedge as long as the true risk premium is non-zero. By ignoring the dynamic effects of learning, utility losses can grow with the investment horizon.

However, the bottom figure reveals that annualized utility losses can both increase and decrease with the investment horizon. For low enough levels of risk aversion or very precise beliefs (not shown on the graph), this will typically not be the case. But for high levels of risk aversion, or if estimation risk is high, the annualized utility loss may actually decrease with the investment horizon.

3.6 Concluding Remarks

In this chapter I have studied the optimal currency hedge for an international investor who is concerned with estimation risk. Increasing estimation risk and investment horizon reduce the speculative demand for holding foreign currency, as well as the market timing motives generated by a time-varying risk premium; the optimal currency hedge is much closer to a complete hedge.

While the hedging policy of a long-run investor looks similar to a complete hedge, an important question deals with the economic importance of estimation risk. Analytical results suggest that, despite the effect of estimation risk on the hedging strategy, the opportunity cost of engaging in a complete hedge grows with the investment horizon, due to the effects of dynamic learning. As posterior beliefs about the risk premium become more precise, foreign currency becomes a more attractive investment. An investor who commits to ignoring the benefits of foreign currency will suffer utility losses today, even when there is a lot of uncertainty about the risk premium.

¹³To convert the opportunity cost into an annuity stream, define the annualized cost LA as

$$LA = \frac{L \times r_d}{1 - \exp\{-r_d \times T\}}$$

3.7 Figures

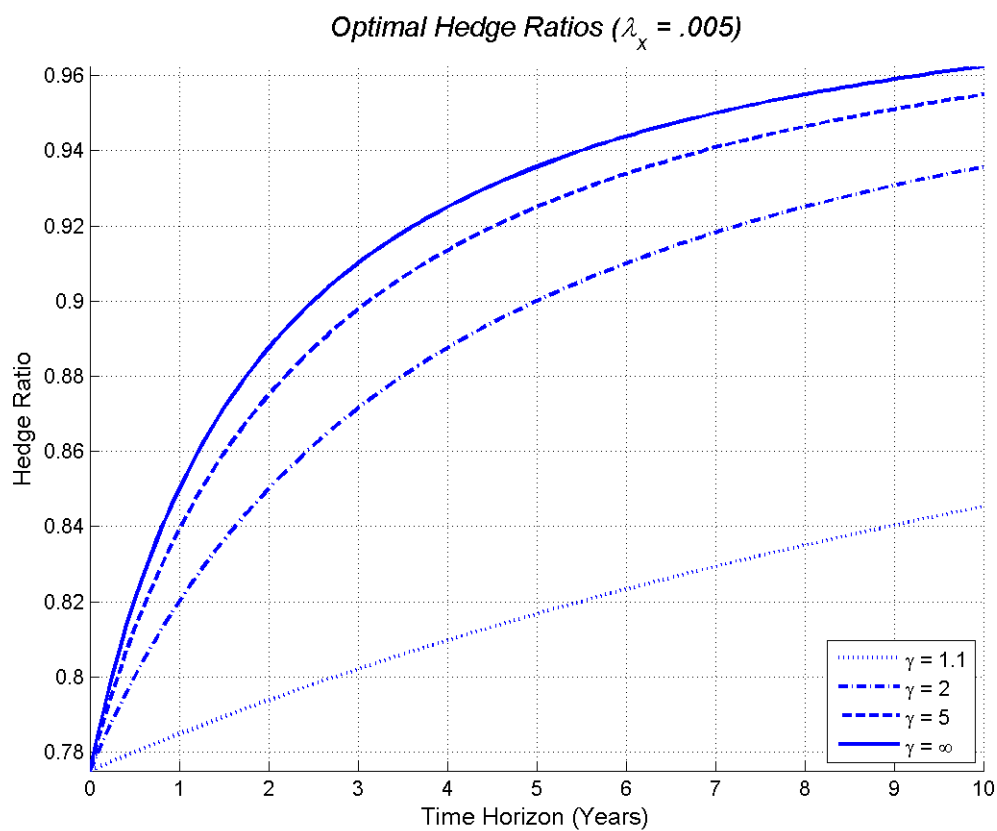


Figure 3.1: Optimal currency hedge ratios as a function of investment horizon, measured in years. The value of the estimated currency risk premium is 0.005, the exchange rate and equity volatilities are 0.10 and 0.15, respectively, and the prior variance is 0.005.

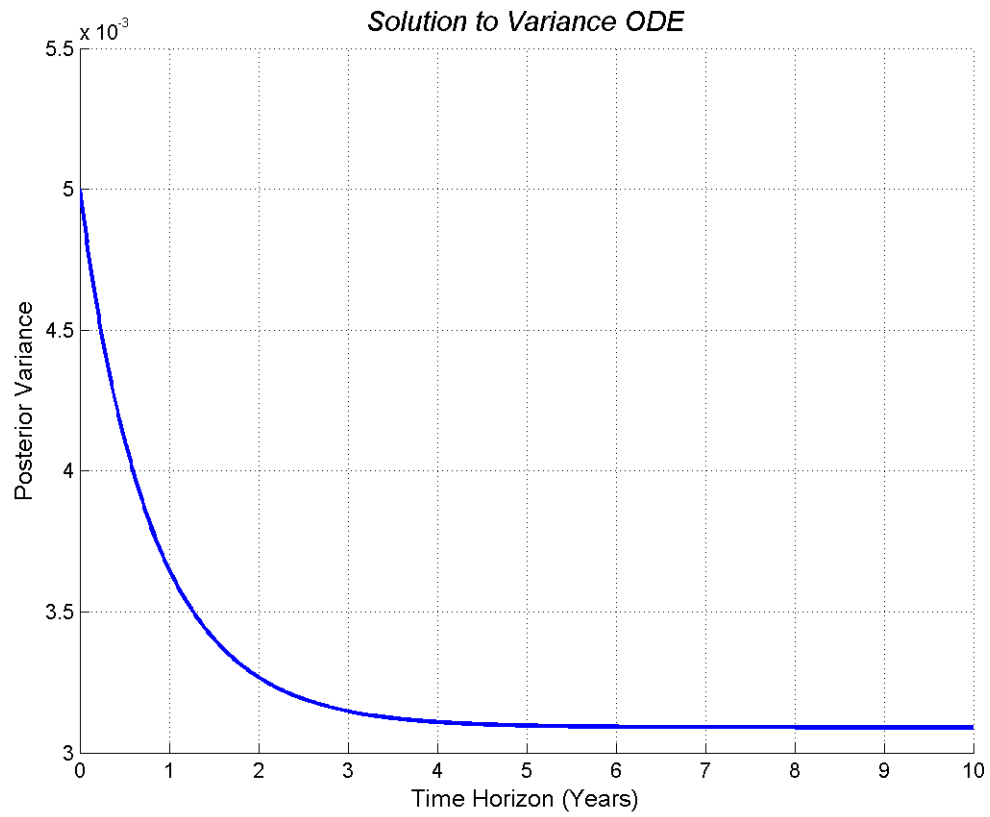


Figure 3.2: Posterior variance as a function of time. The prior estimate is 0.005, the volatilities of the risk premium and exchange rate are 0.05 and 0.10, respectively, and the rate of mean reversion is 0.25. All correlations are zero.

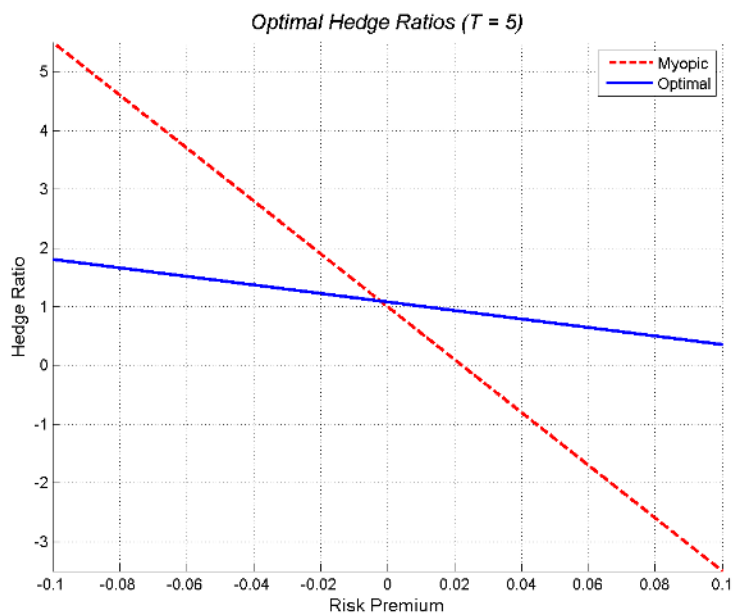


Figure 3.3: Optimal hedge ratio as a function of the estimated risk premium. The prior variance is 0.005, and the volatilities of the exchange rate, exchange rate risk premium, and foreign stock are 0.10, 0.05, and 0.15, respectively. The time horizon is 5 years.

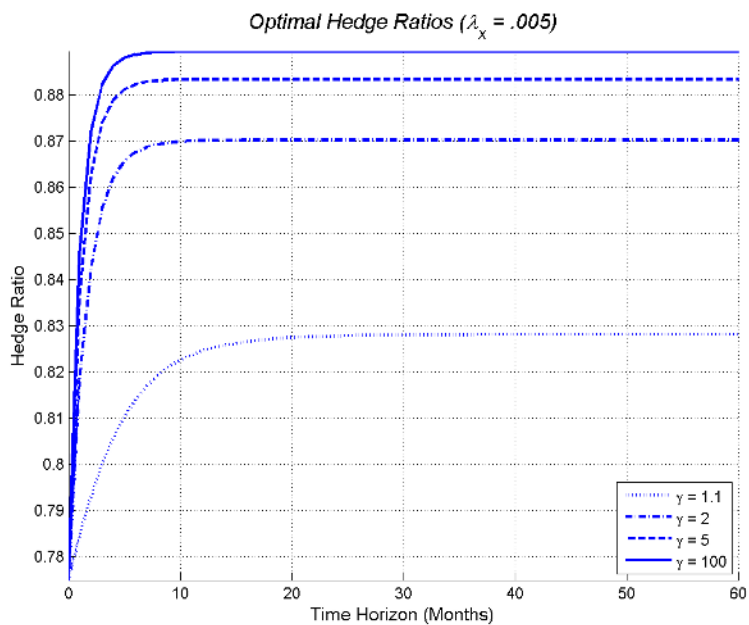


Figure 3.4: Optimal currency hedge ratios as a function of investment horizon, measured in months. The value of the estimated currency risk premium is 0.005, equal to the long-run mean, and the variance of the estimate is 0.005.

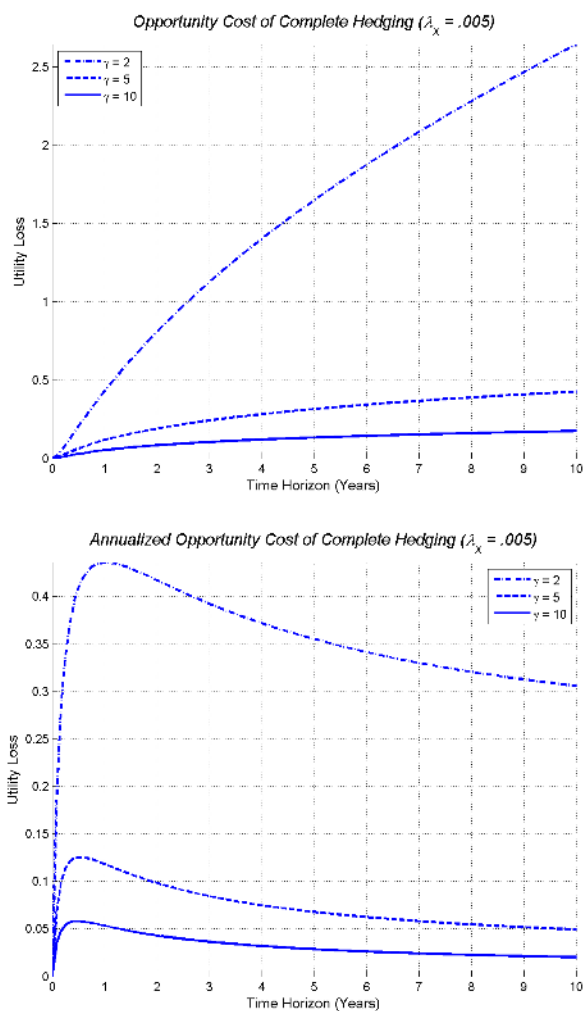


Figure 3.5: Opportunity cost of following a complete currency hedge, measured as the percentage increase in initial wealth necessary to make an investor indifferent between the complete hedging strategy and the optimal hedging strategy, as a function of investment horizon. The top figure presents the total opportunity cost at $t = 0$, while the bottom figure presents the opportunity cost as an annuity stream. The value of the estimated currency risk premium is 0.005, and the variance of the estimate is 0.05.

Bibliography

- Andrew B. Abel and Janice C. Eberly. An exact solution for the investment and value of a firm facing uncertainty, adjustment costs, and irreversibility. *Journal of Economic Dynamics and Control*, 21(4):831–852, 1997.
- T.W. Anderson and Cheng Hsiao. Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18(1):47–82, 1982.
- Manuel Arrelano and Stephen Bond. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58:277–297, 1991.
- David K. Backus, Silverio Foresi, and Chris I. Telmer. Affine term structure models and the forward premium anomaly. *Journal of Finance*, 56:279–304, 2001.
- Ben S. Bernanke. Irreversibility, uncertainty, and cyclical investment. *Quarterly Journal of Economics*, 97(1):141–195, 1983.
- Antonio E. Bernardo and Bhagwan Chowdhry. Resources, real options, and corporate strategy. *Journal of Financial Economics*, 63(2):211–234, 2002.
- Richard Blundell, Stephen Bond, Michael Devereux, and Fabio Schiantarelli. Investment and tobin’s q: Evidence from company panel data. *Journal of Econometrics*, 51(1):233–257, 1992.
- Stephen Bond. Dynamic panel data models: A guide to micro data methods and practice. Centre for Microdata Methods and Practice Working Paper CWP09/02, 2002.
- James R. Brown and Bruce C. Petersen. Why has the investment-cash flow sensitivity declined so sharply? rising r&d and equity market developments. *Journal of Banking and Finance*, 33:971–984, 2009.
- A. Craig Burnside, Martin S. Eichenbaum, Isaac Kleshchelski, and Sergio Rebelo. Do peso problems explain the returns to the carry trade? Working Paper, 2008.

- John Y. Campbell, Andrew W. Lo, and A. Craig MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, 1997.
- John Y. Campbell, Karine Serfaty de Medeiros, and Luis M. Viceira. Global currency hedging. Working Paper, 2008.
- Yu-Fu Chen and Michael Funke. Option value, policy uncertainty, and the foreign direct investment decision. Working Paper, 2003.
- Alexander David. Fluctuating confidence in stock markets: Implications for returns and volatility. *Journal of Financial and Quantitative Analysis*, 32(4):427–462, 1997.
- Alexander David and Pietro Veronesi. Option prices with uncertain fundamentals. Working Paper, 2002.
- Pasquale Della Corte, Lucio Sarno, and Ilias Tsiakas. An economic evaluation of empirical exchange rate models. *Review of Financial Studies*, Forthcoming, 2008.
- Avinash K. Dixit and Robert S. Pindyck. *Investment Under Uncertainty*. Princeton University Press, Princeton, 1994.
- Charles Engel. The forward discount anomaly and the risk premium: A survey of recent evidence. *Journal of Empirical Finance*, 3:123–192, 1996.
- Timothy Erickson and Toni M. Whited. Measurement error and the relationship between investment and q. *Journal of Political Economy*, 108(5):1027–1057, 2000.
- Steve M. Fazzari, Glenn R. Hubbard, and Bruce C. Petersen. Financing constraints and corporate investment. *Brookings Papers on Economic Activity*, 1:141–195, 1988.
- Kenneth A. Froot. Currency hedging over long horizons. NBER Working Paper No. 4355, 1993.
- Rene Garcia. Null distribution of the likelihood ratio test in markov switching models. *International Economic Review*, 39:763–788, 1998.
- Gerard Gennotte. Optimal portfolio choice under incomplete information. *Journal of Finance*, 41:733–746, 1986.
- James D. Hamilton. *Time Series Analysis*. Princeton University Press, Princeton, 1994.
- Kevin A. Hassett and Gilbert E. Metcalf. Investment with uncertain tax policy: Does random tax policy discourage investment? *The Economic Journal*, 109(457): 372–393, 1999.

- Brandon Julio and Youngsuk Yook. Political uncertainty and corporate investment cycles. Working Paper, 2009.
- Steven N. Kaplan and Luigi Zingales. Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics*, 112(1): 169–215, 1997.
- Tong Suk Kim and Edward Omberg. Dynamic nonmyopic portfolio behavior. *Review of Financial Studies*, 9:141–161, 1996.
- Manuel Klein. Irreversible investment under incomplete information. Working Paper, 2007.
- Linda Sandris Larsen and Claus Munk. The costs of suboptimal dynamic asset allocation: General results and applications to interest rate risk, stock volatility risk, and growth/value tilts. Working Paper, 2010.
- Jonathan Lewellen and Katharina Lewellen. Investment and cashflow. Working Paper, 2010.
- Robert S. Liptser and Albert N. Shiryaev. *Statistics of Random Processes I: General Theory*. Springer-Verlag, New York, second edition, 2001.
- Robert S. Liptser and Albert N. Shiryaev. *Statistics of Random Processes II: Applications*. Springer-Verlag, New York, 1978.
- Sanjay K. Nawalkha, Natalie A. Beliaeva, and Gloria M. Soto. *Dynamic Term Structure Modeling: The Fixed Income Valuation Course*. John Wiley and Sons, Hoboken, 2007.
- Christopher J. Neely and Paul A. Weller. Predictability in international asset returns: A reexamination. *Journal of Financial and Quantitative Analysis*, 35:601–620, 2000.
- Robert Novy-Marx. Investment cash-flow sensitivity and the value premium. Working Paper, 2006.
- Dani Rodrik. Policy uncertainty and private investment in developing countries. *Journal of Development Economics*, 36:229–242, 1991.
- David Roodman. How to do xtabond2: An introduction to ‘difference’ and ‘system’ gmm in stata. Center for Global Development Working Paper Number 103, 2006.
- Lucio Sarno. Viewpoint: Towards a solution to the puzzles in exchange rate economics: Where do we stand? *Canadian Journal of Economics*, 38:673–708, 2005.

- Mark E. Schaffer. `xtivreg2`: Stata module to perform extended iv/2sls, gmm, and ac/hac, liml, and k-class regression for panel data models. <http://ideas.repec.org/c/boc/bocode/s456501.html>, 2007.
- Steven E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer, New York, 2004.
- Jeremy J. Siegel. Risk, interest rates, and the forward exchange. *Quarterly Journal of Economics*, 86:303–309, 1972.
- Bruno H. Solnik. An equilibrium model of the international capital market. *Journal of Economic Theory*, 8:500–24, 1974.
- Pablo T. Spiller. Politicians, interest groups, and regulators: A multi-principals agency theory of regulation, or ‘let them be bribed’. *Journal of Law and Economics*, 33(1):65–101, 1990.
- James H. Stock and Mark W. Watson. Heteroskedasticity-robust standard errors for fixed effects panel data regression. *Econometrica*, 76(1):155–174, 2008.
- Pietro Veronesi. How does information quality affect stock returns? *Journal of Finance*, 55(2):807–837, 2000.
- O. Miguel Villanueva. Forecasting currency excess returns: Can we exploit the forward bias? Working Paper, 2007.
- Justin Wolfers and Eric Zitzewitz. Interpreting prediction market prices as probabilities. NBER Working Paper No. 12200, 2006.

Appendix A

Appendix to Chapter 1

A.1 The Filtering Equations

Consider the observable processes Y and s , which follow the system of stochastic differential equations,

$$\frac{dY(t)}{Y(t)} = \mu(\theta, t)dt + \sigma dZ_Y \quad (\text{A.1.1a})$$

$$ds(t) = \theta(t)dt + \eta\rho dZ_Y + \eta\sqrt{1 - \rho^2}dZ_s, \quad (\text{A.1.1b})$$

where Z_Y and Z_s are standard, orthogonal Brownian motions. $\theta(t) \in \{\theta_l, \theta_h\}$ is a state variable controlling the drifts of Y and s and follows a two-state Markov process with transition matrix

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_{ll} & \lambda_{lh} \\ \lambda_{hl} & \lambda_{hh} \end{pmatrix} = \begin{pmatrix} -\lambda_{lh} & \lambda_{lh} \\ \lambda_{hl} & -\lambda_{hl} \end{pmatrix}. \quad (\text{A.1.2})$$

Optimal filtering equations make use of two key assumptions:

Assumption 1. *The signals Y and s are observable, but the Brownian motions Z_Y and Z_s are unobservable.*

Assumption 2. *θ follows a hidden Markov process. The parameters of the process $(\theta_l, \theta_h, \mathbf{\Lambda})$ are known, but the current state $\theta(t)$ is unknown.*

The filtering problem boils down to finding estimates for $\theta(t)$ and $\mu(\theta, t)$ based on observations of Y and s . The optimal estimates minimize mean-squared error, and are given by

$$\hat{\theta}(\pi, t) = E[\theta(t)|\pi(t)] = \pi(t)\theta_h + (1 - \pi(t))\theta_l \quad (\text{A.1.3a})$$

$$\hat{\mu}(\pi, t) = E[\mu(\theta, t)|\pi(t)] = \pi(t)\mu(\theta_h) + (1 - \pi(t))\mu(\theta_l) \quad (\text{A.1.3b})$$

where

$$\pi(t) = \text{Prob}(\theta(t) = \theta_h | \mathcal{F}(t)) \quad (\text{A.1.4})$$

is the posterior probability. Formally, $\mathcal{F}(t)$ is the filtration generated by Y and s , which describes how information is revealed over time.

Case 1 - One Signal

Suppose that only the signal Y is observed. Optimal filtering equations follow from Theorem 9.1 of Liptser and Shiryaev (2001), which I state as the following Lemma:

Lemma 1. *Given Assumptions 1 and 2, the posterior belief π satisfies the stochastic differential equation*

$$d\pi(t) = \kappa(\bar{\pi} - \pi(t))dt + \pi(t)(1 - \pi(t)) \frac{\mu(\theta_h) - \mu(\theta_l)}{\sigma^2} \left[\frac{dY(t)}{Y(t)} - \hat{\mu}(\pi, t)dt \right] \quad (\text{A.1.5})$$

with

$$\begin{aligned} \kappa &= \lambda_{lh} + \lambda_{hl} \\ \bar{\pi} &= \frac{\lambda_{lh}}{\lambda_{lh} + \lambda_{hl}}. \end{aligned}$$

While equation (A.1.5) fully characterizes the learning process, the stochastic equations for Y and π are still written in terms of unobservable parameters and shocks, specifically $\theta(t)$ and $dZ_Y(t)$. The next step is to transform the incomplete information system into a complete information system. To do so, rewrite the bracketed expression in (A.1.5) as

$$\begin{aligned} \frac{dY(t)}{Y(t)} - \hat{\mu}(\pi, t) &= [\mu(\theta, t) - \hat{\mu}(\pi, t)]dt + \sigma dZ_Y(t) \\ &= \sigma \left[\frac{\mu(\theta, t) - \hat{\mu}(\pi, t)}{\sigma} dt + dZ_Y(t) \right] \\ &= \sigma d\hat{Z}_Y(t) \end{aligned}$$

where

$$\hat{Z}_Y(t) = \frac{\mu(\theta, t) - \hat{\mu}(\pi, t)}{\sigma} t + Z_Y(t)$$

is a Brownian motion adapted to the filtration $\mathcal{F}(t)$, i.e. an observable Brownian motion. Substitution of $d\hat{Z}_Y$ into the processes for Y and π delivers the complete information system of stochastic equations.

Corollary 1. *The incomplete information system defined in (A.1.1), (A.1.2), and (A.1.5), can be characterized by the complete information system*

$$\frac{dY(t)}{Y(t)} = \hat{\mu}(\pi, t)dt + \sigma d\hat{Z}_Y(t) \quad (\text{A.1.7a})$$

$$d\pi(t) = \kappa(\bar{\pi} - \pi(t))dt + \pi(t)(1 - \pi(t))\frac{\mu(\theta_h) - \mu(\theta_l)}{\sigma}d\hat{Z}_Y(t) \quad (\text{A.1.7b})$$

This is a complete information system because the stochastic processes for Y and π are written in terms of the observable parameter $\hat{\mu}(\pi, t)$ and the observable Brownian motion \hat{Z}_Y .

Finally, an application of Ito's Lemma $\hat{\mu}$ gives the stochastic differential equation for the conditional expectation of μ ,

$$d\hat{\mu}(\pi, t) = \kappa(\bar{\mu} - \hat{\mu}(\pi, t))dt + \pi(t)(1 - \pi(t))\frac{(\mu(\theta_h) - \mu(\theta_l))^2}{\sigma}d\hat{Z}_Y \quad (\text{A.1.8})$$

with

$$\bar{\mu} = \frac{\lambda_{lh}}{\lambda_{lh} + \lambda_{hl}}\mu(\theta_h) + \frac{\lambda_{hl}}{\lambda_{lh} + \lambda_{hl}}\mu(\theta_l) \quad (\text{A.1.9})$$

Case 2 - Two Signals

Now suppose that both signals Y and s are observed. To derive the stochastic process for π in the two-variable case, it is convenient to stack the observable signals into a vector \mathbf{S} :

$$\begin{bmatrix} dY(t) \\ ds(t) \end{bmatrix} = \begin{bmatrix} \mu(\theta, t)Y(t) \\ \theta(t) \end{bmatrix} dt + \begin{bmatrix} \sigma Y(t) & 0 \\ \eta\rho & \eta\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} dZ_Y(t) \\ dZ_s(t) \end{bmatrix}.$$

Or, in vector notation,

$$d\mathbf{S}(t) = \begin{bmatrix} dY(t) \\ ds(t) \end{bmatrix} = \mathbf{\Gamma}(\theta, \mathbf{S})dt + \mathbf{\Sigma}(\mathbf{S})d\mathbf{Z}_S(t). \quad (\text{A.1.10})$$

Lemma 1 from Veronesi (2000) generalizes Theorem 9.1 from Liptser and Shiryaev (2001) to multiple signals.

Lemma 2. *Given Assumptions 1 and 2, the posterior belief π satisfies the stochastic*

process

$$\begin{aligned}
d\pi(t) = & \kappa(\bar{\pi} - \pi(t)) dt + \\
& \left(\pi(t)(1 - \pi(t)) \right. \\
& \times \left. (\mathbf{\Gamma}(\theta_h, \mathbf{S}) - \mathbf{\Gamma}(\theta_l, \mathbf{S}))' (\mathbf{\Sigma}(\mathbf{S})\mathbf{\Sigma}(\mathbf{S})')^{-1} \left[d\mathbf{S}(t) - \hat{\mathbf{\Gamma}}(\pi, \mathbf{S}) dt \right] \right)
\end{aligned} \tag{A.1.11}$$

with

$$\begin{aligned}
\hat{\mathbf{\Gamma}}(\pi, \mathbf{S}) &= \pi(t)\mathbf{\Gamma}(\theta_h, \mathbf{S}) + (1 - \pi(t))\mathbf{\Gamma}(\theta_l, \mathbf{S}) \\
\kappa &= \lambda_{lh} + \lambda_{hl} \\
\bar{\pi} &= \frac{\lambda_{lh}}{\lambda_{lh} + \lambda_{hl}}.
\end{aligned}$$

Written in scalar form, equation (A.1.11) is

$$\begin{aligned}
d\pi(t) = & \kappa(\bar{\pi} - \pi(t)) dt \\
& + \pi(t)(1 - \pi(t)) \left(\frac{\mu(\theta_h) - \mu(\theta_l)}{\sigma^2(1 - \rho^2)} - \frac{\theta_h - \theta_l}{\eta\sigma} \frac{\rho}{1 - \rho^2} \right) \left(\frac{dY(t)}{Y(t)} - \hat{\mu}(\pi, t) \right) \\
& + \pi(t)(1 - \pi(t)) \left(\frac{\theta_h - \theta_l}{\eta^2(1 - \rho^2)} - \frac{\mu(\theta_h) - \mu(\theta_l)}{\eta\sigma} \frac{\rho}{1 - \rho^2} \right) \left(ds(t) - \hat{\theta}(\pi, t) \right)
\end{aligned} \tag{A.1.13}$$

As with the scalar case, by making the appropriate substitutions, the stochastic differential equations describing the evolution of \mathbf{S} and π have a complete information representation. Again, begin by rewriting the bracketed expression in (A.1.11) as

$$\begin{aligned}
d\mathbf{S}(t) - \hat{\mathbf{\Gamma}}(\pi, \mathbf{S}) dt &= \left(\mathbf{\Gamma}(\theta, \mathbf{S}) - \hat{\mathbf{\Gamma}}(\pi, \mathbf{S}) \right) dt + \mathbf{\Sigma}(\mathbf{S}) d\mathbf{Z}_{\mathbf{S}}(t) \\
&= \mathbf{\Sigma}(\mathbf{S}) \left[\mathbf{\Sigma}(\mathbf{S})^{-1} \left(\mathbf{\Gamma}(\theta, \mathbf{S}) - \hat{\mathbf{\Gamma}}(\pi, \mathbf{S}) \right) dt + d\mathbf{Z}_{\mathbf{S}}(t) \right] \\
&= \mathbf{\Sigma}(\mathbf{S}) d\hat{\mathbf{Z}}_{\mathbf{S}}(t)
\end{aligned}$$

where

$$\hat{\mathbf{Z}}_{\mathbf{S}}(t) = \mathbf{\Sigma}(\mathbf{S})^{-1} \left(\mathbf{\Gamma}(\theta, \mathbf{S}) - \hat{\mathbf{\Gamma}}(\pi, \mathbf{S}) \right) t + \mathbf{Z}_{\mathbf{S}}(t)$$

is a two-dimensional vector Brownian motion adapted to the filtration $\mathcal{F}(t)$, i.e. an observable Brownian motion. Following the same steps as in the scalar case, substitution of $d\hat{\mathbf{Z}}_{\mathbf{S}}(t)$ into the processes for \mathbf{S} and π delivers the complete information system of stochastic differential equations, where the signal vector \mathbf{S} and posterior belief π are written in terms of the observable processes $\hat{\mathbf{\Gamma}}$ and $\hat{\mathbf{Z}}_{\mathbf{S}}$.

Corollary 2. *The incomplete information system defined in (A.1.2), (A.1.10), and (A.1.11), can be characterized by the complete information system*

$$d\mathbf{S}(t) = \hat{\Gamma}(\pi, \mathbf{S})dt + \Sigma(\mathbf{S})d\hat{\mathbf{Z}}_{\mathbf{S}}(t) \quad (\text{A.1.14a})$$

$$\begin{aligned} d\pi(t) &= \kappa(\bar{\pi} - \pi(t))dt + \pi(t)(1 - \pi(t)) \\ &\quad \times [\Sigma(\mathbf{S})^{-1}(\Gamma(\theta_h, \mathbf{S}) - \Gamma(\theta_l, \mathbf{S}))]' d\hat{\mathbf{Z}}_{\mathbf{S}}(t). \end{aligned} \quad (\text{A.1.14b})$$

While the equations in (A.1.14) fully characterize the complete information system, it is helpful to have the equations for Y , s , and π in scalar form. From the definition of $\hat{\mathbf{Z}}_{\mathbf{S}}$,

$$\begin{aligned} d\hat{\mathbf{Z}}_{\mathbf{S}}(t) &= \Sigma(\mathbf{S})^{-1} \left(\Gamma(\theta, \mathbf{S}) - \hat{\Gamma}(\pi, \mathbf{S}) \right) dt + d\mathbf{Z}_{\mathbf{S}}(t) \\ &= \left[\begin{array}{c} \frac{\mu(\theta, t) - \hat{\mu}(\pi, t)}{\eta\sqrt{1-\rho^2}} - \frac{\mu(\theta, t) - \hat{\mu}(\pi, t)}{\sigma} \frac{\rho}{\sqrt{1-\rho^2}} \end{array} \right] dt + \left[\begin{array}{c} dZ_Y(t) \\ dZ_s(t) \end{array} \right] \\ &= \left[\begin{array}{c} d\hat{Z}_Y(t) \\ d\hat{Z}_s(t) \end{array} \right]. \end{aligned}$$

Then the processes for Y and s are

$$\frac{dY(t)}{Y(t)} = \hat{\mu}(\pi, t)dt + \sigma d\hat{Z}_Y(t) \quad (\text{A.1.15a})$$

$$ds(t) = \hat{\theta}(\pi, t)dt + \eta \left(\rho d\hat{Z}_Y(t) + \sqrt{1-\rho^2} d\hat{Z}_s(t) \right). \quad (\text{A.1.15b})$$

where \hat{Z}_Y and \hat{Z}_s are orthogonal Brownian motions adapted to $\mathcal{F}(t)$. Next, define $\Sigma_{\pi} = \Sigma(\mathbf{S})^{-1}(\Gamma(\theta_h, \mathbf{S}) - \Gamma(\theta_l, \mathbf{S}))$. Then multiplying through,

$$\Sigma_{\pi} = \left[\begin{array}{c} \frac{\mu(\theta, t) - \hat{\mu}(\pi, t)}{\eta\sqrt{1-\rho^2}} - \frac{\mu(\theta, t) - \hat{\mu}(\pi, t)}{\sigma} \frac{\rho}{\sqrt{1-\rho^2}} \end{array} \right]$$

Finally, using Σ_{π} , re-write the stochastic process for π in scalar form,

$$\begin{aligned} d\pi(t) &= \kappa(\bar{\pi} - \pi(t))dt + \pi(t)(1 - \pi(t))\Sigma'_{\pi}d\hat{\mathbf{Z}}_{\mathbf{S}}(t) \\ &= \kappa(\bar{\pi} - \pi(t))dt + \pi(t)(1 - \pi(t)) \left(\frac{\mu(\theta_h) - \mu(\theta_l)}{\sigma} \right) d\hat{Z}_Y(t) \\ &\quad + \pi(t)(1 - \pi(t)) \left(\frac{\theta_h - \theta_l}{\eta\sqrt{1-\rho^2}} - \frac{\mu(\theta_h) - \mu(\theta_l)}{\sigma} \frac{\rho}{\sqrt{1-\rho^2}} \right) d\hat{Z}_s(t) \end{aligned} \quad (\text{A.1.16})$$

A.2 Properties of the Learning Process

A.2.1 Conditional Moments

Two features of the learning process are worth pointing out. The first is that beliefs mean revert to $\bar{\pi}$ at rate κ . Indeed, given a prior $\pi(0)$, the conditional expectation of the posterior probability is given by

$$\begin{aligned} E[\pi(t)|\pi(0)] &= \pi(0)e^{-\kappa t} + \bar{\pi}(1 - e^{-\kappa t}) \\ &= \pi(0)e^{-(\lambda_{lh} + \lambda_{hl})t} + \left(\frac{\lambda_{lh}}{\lambda_{lh} + \lambda_{hl}}\right) \left(1 - e^{-(\lambda_{lh} + \lambda_{hl})t}\right) \end{aligned} \quad (\text{A.2.1})$$

with

$$\lim_{t \rightarrow \infty} E[\pi(t)|\pi(0)] = \bar{\pi},$$

a weighted average of the prior and the long-run mean. This is the familiar form for mean-reverting stochastic processes.

To derive this expectation, first write the stochastic differential equation for π in integral form,

$$\pi(t) - \pi(0) = \int_0^t \kappa(\bar{\pi} - \pi(u)) du + \int_0^t \pi(u)(1 - \pi(u)) \Sigma_{\pi} d\hat{\mathbf{Z}}_{\mathcal{S}}(u). \quad (\text{A.2.2})$$

Next, define $m(t) = E[\pi(t)|\pi(0)]$ and take expectations, noting that the expectation of a stochastic integral is zero:

$$\begin{aligned} m(t) &= E[\pi(t)|\pi(0)] = \pi(0) + \int_0^t \kappa(\bar{\pi} - E[\pi(u)|\pi(0)]) du \\ &= \pi(0) + \kappa\bar{\pi}t - \kappa \int_0^t m(u) du. \end{aligned}$$

Taking derivatives yields an ordinary differential equation describing the evolution of the conditional mean,

$$m'(t) = \kappa\bar{\pi} - \kappa m(t),$$

with general solution

$$m(t) = \bar{\pi} + e^{-\kappa t} C, \quad (\text{A.2.3})$$

where C is an arbitrary constant of integration. The initial condition $m(0) = \pi(0)$ pins down $C = \pi(0) - \bar{\pi}$; substitution into (A.2.3) delivers the solution in (A.2.1).

The second property is that one statistic, π , characterizes both means and vari-

ances, which are given by

$$\hat{\theta}(\pi) = E[\theta|\pi] = \pi\theta_1 + (1 - \pi)\theta_0 \quad (\text{A.2.4a})$$

$$\hat{\mu}(\pi, X) = E[\mu(\theta, X)|\pi] = \pi\mu(\theta_1, X) + (1 - \pi)\mu(\theta_0, X) \quad (\text{A.2.4b})$$

and

$$\text{Var}(\theta) = E[(\theta - \hat{\theta})^2|\pi] = \pi(1 - \pi)(\theta_h - \theta_l)^2 \quad (\text{A.2.5a})$$

$$\text{Var}(\mu(\theta, X)) = E[(\mu(\theta, X) - \hat{\mu})^2|\pi] = \pi(1 - \pi)(\mu(\theta_h, X) - \mu(\theta_l, X))^2. \quad (\text{A.2.5b})$$

Intuition for this result comes from the diffusion term in equation (A.1.7). When π is close to zero or one, the diffusion term, which is proportional to $\pi(1 - \pi)$, is small; beliefs, and therefore estimates of θ and $\mu(\theta, X)$, are precise. However, when beliefs are close to $\pi = \frac{1}{2}$, there is a large amount of uncertainty about the state variable, and estimates are relatively imprecise.

A.3 Bellman Equation

Conditional on an amount of invested capital K , the present value of future cash flows at time t is given by equation (1.19) in Section 1.2,

$$V(Y, \pi; K, t) = E_t \left[\int_t^\infty e^{-r(u-t)} Y(u) F(K)^{1-\frac{1}{\gamma}} du \right]. \quad (\text{A.3.1})$$

The value of the cash flows, $V(Y, \pi; K, t)$ satisfies the Bellman equation given in (1.20),

$$rV dt = YF(K)^{1-\frac{1}{\gamma}} dt + E[dV], \quad (\text{A.3.2})$$

which relates the total return from a perpetual stream of cash flows (dividend yield plus capital gain) to the risk-free return. To derive this relationship, rewrite (A.3.1) as the sum of two components,

$$\begin{aligned} V(Y, \pi; K, t) &= E_t \left[\int_t^\tau e^{-r(u-t)} Y(u) F(K)^{1-\frac{1}{\gamma}} du \right. \\ &\quad \left. + e^{-r(\tau-t)} \int_\tau^\infty e^{-r(u-\tau)} Y(u) F(K)^{1-\frac{1}{\gamma}} du \right] \\ &= E_t \left[\int_t^\tau e^{-r(u-t)} Y(u) F(K)^{1-\frac{1}{\gamma}} du + e^{-r(\tau-t)} V(Y, \pi; K, \tau) \right]. \end{aligned} \quad (\text{A.3.3})$$

Subtracting $V(Y, \pi; K, t)$ from both sides gives

$$0 = E_t \left[\int_t^\tau e^{-r(u-t)} Y(u) F(K)^{1-\frac{1}{\gamma}} du + e^{-r(\tau-t)} V(Y, \pi; K, \tau) - V(Y, \pi; K, t) \right]. \quad (\text{A.3.4})$$

Next, define $X(t) = e^{-rt} V(Y, \pi; K, t)$. Substitution into (A.3.4) gives

$$0 = E_t \left[\int_t^\tau e^{-r(u-t)} Y(u) F(K)^{1-\frac{1}{\gamma}} du + e^{rt} (X(\tau) - X(t)) \right]. \quad (\text{A.3.5})$$

Taking limits,

$$\begin{aligned} 0 &= \lim_{\tau \rightarrow t} E_t \left[\int_t^\tau e^{-r(u-t)} Y(u) F(K)^{1-\frac{1}{\gamma}} du + e^{rt} (X(\tau) - X(t)) \right] \\ &= Y(t) F(K)^{1-\frac{1}{\gamma}} dt + e^{rt} E_t [dX]. \end{aligned} \quad (\text{A.3.6})$$

Applying Ito's Lemma to X gives $dX = e^{-rt} dV - r e^{-rt} V dt$. Direct substitution into (A.3.6) delivers the Bellman equation in (A.3.2).

A.4 Maximum Likelihood Estimation

The parameter vector is

$$\Theta = [\mu_h \quad \mu_l \quad \sigma \quad \lambda_{hl} \quad \lambda_{lh}]' \quad (\text{A.4.1})$$

and the log-likelihood function is

$$\mathcal{L}(\Theta) = \sum_{t=1}^T \log f(y(t)) \quad (\text{A.4.2})$$

with

$$\begin{aligned} f(y(t)) &= (1 - \lambda_{hl}) \pi_{t-1} f_h + \lambda_{hl} \pi_{t-1} f_l \\ &\quad + \lambda_{lh} (1 - \pi_{t-1}) f_h + (1 - \lambda_{lh}) (1 - \pi_{t-1}) f_l. \end{aligned} \quad (\text{A.4.3})$$

and

$$f_h = f(y(t)|\theta(t) = \theta_h) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y(t) - \mu_h + \frac{1}{2}\sigma^2}{\sigma} \right)^2 \right\} \quad (\text{A.4.4a})$$

$$f_l = f(y(t)|\theta(t) = \theta_l) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y(t) - \mu_l + \frac{1}{2}\sigma^2}{\sigma} \right)^2 \right\}. \quad (\text{A.4.4b})$$

Calculating standard errors for the parameter estimates relies on a sample estimate of the information matrix, given by

$$\hat{I} = \frac{1}{T} \sum_{t=1}^T \frac{\partial \log f(\hat{\Theta})}{\partial \Theta} \frac{\partial \log f(\hat{\Theta})'}{\partial \Theta} \quad (\text{A.4.5})$$

The relevant derivatives are

$$\begin{aligned} \frac{\partial \log f}{\partial \mu_h} &= \frac{(1 - \lambda_{hl})\pi_{t-1} + \lambda_{lh}(1 - \pi_{t-1})}{f(y(t))} \frac{\partial f_h}{\partial \mu_h} \\ \frac{\partial \log f}{\partial \mu_l} &= \frac{(1 - \lambda_{lh})(1 - \pi_{t-1}) + \lambda_{hl}\pi_{t-1}}{f(y(t))} \frac{\partial f_l}{\partial \mu_l} \\ \frac{\partial \log f}{\partial \sigma} &= \frac{1}{f(y(t))} \left[((1 - \lambda_{hl})\pi_{t-1} + \lambda_{lh}(1 - \pi_{t-1})) \frac{\partial f_h}{\partial \sigma} \right. \\ &\quad \left. + ((1 - \lambda_{lh})(1 - \pi_{t-1}) + \lambda_{hl}\pi_{t-1}) \frac{\partial f_l}{\partial \sigma} \right] \\ \frac{\partial \log f}{\partial \lambda_{hl}} &= \frac{\pi_{t-1}(f_h - f_l)}{f(y(t))} \\ \frac{\partial \log f}{\partial \lambda_{lh}} &= \frac{(1 - \pi_{t-1})(f_l - f_h)}{f(y(t))} \end{aligned} \quad (\text{A.4.6})$$

with

$$\begin{aligned} \frac{\partial f_h}{\partial \mu_h} &= f_h \left(\frac{y(t) - \mu_h + \frac{1}{2}\sigma^2}{\sigma^2} \right) \\ \frac{\partial f_l}{\partial \mu_l} &= f_l \left(\frac{y(t) - \mu_l + \frac{1}{2}\sigma^2}{\sigma^2} \right) \\ \frac{\partial f_h}{\partial \sigma} &= f_h \left(\frac{(y(t) - \mu_h)^2}{\sigma^3} - \frac{\sigma}{4} - \frac{1}{\sigma} \right) \\ \frac{\partial f_l}{\partial \sigma} &= f_l \left(\frac{(y(t) - \mu_l)^2}{\sigma^3} - \frac{\sigma}{4} - \frac{1}{\sigma} \right). \end{aligned} \quad (\text{A.4.7})$$

A.5 Numerical Solution to the PDE

Due to the free boundary problem, the partial differential equation in (1.27) requires a numerical solution. I approximate the solution using a finite difference scheme. To solve the PDE, it is convenient to assume the investment option is finite, but with a very long horizon (i.e. 40 years). The finite-horizon PDE is,

$$\frac{\partial G}{\partial t} + L_Y G + L_\pi G + L_{Y\pi} G = 0 \quad (\text{A.5.1})$$

where

$$L_Y G = [\pi\mu_h + (1 - \pi)\mu_l] Y G_Y + \frac{1}{2} \sigma^2 Y^2 G_{YY} - rG \quad (\text{A.5.2a})$$

$$L_\pi G = \kappa(\bar{\pi} - \pi) G_\pi + \frac{1}{2} \pi^2 (1 - \pi)^2 (\omega_Y^2 + \omega_s^2) G_{\pi\pi} \quad (\text{A.5.2b})$$

$$L_{Y\pi} G = \pi(1 - \pi) (\mu_h - \mu_l) Y G_{Y\pi}. \quad (\text{A.5.2c})$$

The next step is to replace L_Y and L_π with implicit finite difference approximations, and $L_{Y\pi}$ with an explicit finite difference approximation:

$$\frac{G_{i,j,k+1} - G_{i,j,k}}{\Delta t} + L_Y G_{i,j,k} + L_\pi G_{i,j,k} + L_{Y\pi} G_{i,j,k+1} \quad (\text{A.5.3})$$

where i indexes Y , j indexes π , and k indexes time. Rearranging,

$$(1 - \Delta t [L_Y + L_\pi]) G_{i,j,k} = (1 + \Delta t L_{Y\pi}) G_{i,j,k+1} \quad (\text{A.5.4})$$

While this system of equations is sparse, it is not tri-diagonal; numerical analysis is possible, though inefficient for small ΔY and $\Delta \pi$. However, using

$$(1 - \Delta t L_Y)(1 - \Delta t L_\pi) = 1 - \Delta t [L_Y + L_\pi] + O(\Delta t^2) \quad (\text{A.5.5})$$

we can approximate (A.5.4) with

$$(1 - \Delta t L_Y)(1 - \Delta t L_\pi) G_{i,j,k} = (1 + \Delta t L_{Y\pi}) G_{i,j,k+1} \quad (\text{A.5.6})$$

which is tri-diagonal. We proceed in two steps:

$$(1 - \Delta t L_Y) G_{i,j,k+\frac{1}{2}} = (1 + \Delta t L_{Y\pi}) G_{i,j,k+1} \quad (\text{A.5.7a})$$

$$(1 - \Delta t L_\pi) G_{i,j,k} = G_{i,j,k+\frac{1}{2}} \quad (\text{A.5.7b})$$

where each step contains a tri-diagonal matrix on the left-hand side, and known values on the right-hand side. This method is known as the ‘operator-splitting’ method and is unconditionally stable. For the grid, I set $Y_{\min} = 0$ with $G(0, \pi) = 0$, and

$Y_{\max} = 1$ with $G(1, \pi) = V(1, \pi; K^*) - K^*$. The lower and upper boundaries for π are 0 and 1. Because $G_{\pi\pi}$ and $G_{Y\pi}$ cancel from the PDE at $\pi = 0$ and $\pi = 1$, I impose no boundary conditions for π ; instead, I approximate L_π with a one-sided finite difference approximation at the boundaries for π .

Appendix B

Appendix to Chapter 3

B.1 ODE for the Posterior Variance

In the general model of portfolio choice presented in Chapter 3, the conditional variance of the estimated currency risk premium, $\nu(t)$, satisfies the ordinary differential equation

$$\frac{d\nu(t)}{dt} = \omega_\lambda^2 - 2 \left(\kappa_\lambda + \frac{\rho_{X\lambda} - \rho_{XS}\rho_{S\lambda}}{1 - \rho_{XS}^2} \frac{\sigma_\lambda}{\sigma_X} \right) \nu(t) - \left(\frac{1}{1 - \rho_{XS}^2} \frac{1}{\sigma_X^2} \right) \nu(t)^2 \quad (\text{B.1.1})$$

with initial condition $\nu(0) = \nu_0$. This is a Riccati differential equation with constant coefficients, and the solution is given in Nawalkha et al. (2007).

To solve the differential equation, rewrite the equation as

$$\frac{d\nu(t)}{dt} = P + Q\nu(t) + R\nu(t)^2 \quad (\text{B.1.2})$$

with

$$P = \omega_\lambda^2 \quad (\text{B.1.3a})$$

$$Q = -2 \left(\kappa_\lambda + \frac{\rho_{X\lambda} - \rho_{XS}\rho_{S\lambda}}{1 - \rho_{XS}^2} \frac{\sigma_\lambda}{\sigma_X} \right) \quad (\text{B.1.3b})$$

$$R = - \left(\frac{1}{1 - \rho_{XS}^2} \frac{1}{\sigma_X^2} \right). \quad (\text{B.1.3c})$$

Then the solution is

$$\nu(t) = -\frac{b}{R} \left[\frac{\frac{a}{b} \left(\frac{b+R\nu_0}{a+R\nu_0} \right) e^{(a-b)t} - 1}{\left(\frac{b+R\nu_0}{a+R\nu_0} \right) e^{(a-b)t} - 1} \right] \quad (\text{B.1.4})$$

with

$$\begin{aligned} a &= \frac{Q + \sqrt{Q^2 - 4PR}}{2} \\ b &= \frac{Q - \sqrt{Q^2 - 4PR}}{2}. \end{aligned} \quad (\text{B.1.5})$$

To find the steady-state value of $\nu(t)$, take the limit as t approaches ∞ :

$$\lim_{t \rightarrow \infty} \nu(t) = -\frac{a}{R} = \frac{Q + \sqrt{Q^2 - 4PR}}{2R} \quad (\text{B.1.6})$$

Notice that the long-run variance is also one of the roots of the quadratic equation that results when setting $d\nu(t)/dt = P + Q\nu(t) + R\nu(t)^2 = 0$:

$$\begin{aligned} f(x) &= P + Qx + Rx^2 = 0 \\ \Rightarrow x &= \frac{-Q \pm \sqrt{Q^2 - 4PR}}{2R} \\ &= \left\{ -\frac{b}{R}, -\frac{a}{R} \right\} \end{aligned} \quad (\text{B.1.7})$$

B.2 Value Function in the General Model

The Bellman equation for the general portfolio choice, given by equation (3.31), is

$$\begin{aligned} 0 = \max_{\alpha} \left\{ \frac{\partial J}{\partial t} + J_W W [r_d + \alpha' \Lambda_P] + \frac{1}{2} J_{WW} W^2 \alpha' \Sigma_P \Sigma_P' \alpha + \kappa_\lambda \left(\frac{1}{2} \sigma_X^2 - \hat{\lambda}_{Xt} \right) J_{\hat{\lambda}} \right. \\ \left. + \frac{1}{2} (\Sigma_\lambda + \nu(t) \Sigma_{\mathcal{I}}^{-1} \mathbf{M}_1)' (\Sigma_\lambda + \nu(t) \Sigma_{\mathcal{I}}^{-1} \mathbf{M}_1) J_{\hat{\lambda}\hat{\lambda}} \right. \\ \left. + W \alpha' \Sigma_P (\Sigma_\lambda + \nu(t) \Sigma_{\mathcal{I}}^{-1} \mathbf{M}_1) J_{W\hat{\lambda}} \right\} \end{aligned} \quad (\text{B.2.1})$$

and the associated optimal portfolio rule is

$$\begin{aligned} \alpha &= -\frac{J_W}{J_{WW} W} (\Sigma_P \Sigma_P')^{-1} \Lambda_P - \frac{J_{W\hat{\lambda}}}{J_{WW} W} (\Sigma_P')^{-1} \Sigma_\lambda \\ &\quad - \frac{J_{W\hat{\lambda}}}{J_{WW} W} (\Sigma_P')^{-1} \Sigma_{\mathcal{I}}^{-1} \mathbf{M}_1 \nu(t). \end{aligned} \quad (\text{B.2.2})$$

To solve the problem, conjecture that the value function $J(W, \hat{\lambda}_{Xt}, t)$ takes the

form

$$J(W, \hat{\lambda}_{Xt}, t) = \frac{W_t^{1-\gamma}}{1-\gamma} F[\hat{\lambda}_{Xt}, t]^\gamma \quad (\text{B.2.3})$$

so that the optimal portfolio rule is given by

$$\boldsymbol{\alpha} = \frac{1}{\gamma} (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P)^{-1} \boldsymbol{\Lambda}_P + \frac{F_\lambda}{F} (\boldsymbol{\Sigma}'_P)^{-1} \boldsymbol{\Sigma}_\lambda + \frac{F_\lambda}{F} (\boldsymbol{\Sigma}'_P)^{-1} \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \mathbf{M}_1 \nu(t). \quad (\text{B.2.4})$$

Substituting the conjecture and the portfolio rule into the Bellman equation and simplifying leads to a partial differential equation for $F(\hat{\lambda}_{Xt}, t)$,

$$\begin{aligned} 0 = & \left\{ \frac{\partial F}{\partial t} + \left[\left(\frac{1}{\gamma} - 1 \right) \left(r_d + \frac{1}{2\gamma} \boldsymbol{\Lambda}'_P (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P)^{-1} \boldsymbol{\Lambda}_P \right) \right] F(\hat{\lambda}_{Xt}, t) \right. \\ & + \left[\kappa_\lambda \left(\frac{1}{2} \sigma_X^2 - \hat{\lambda}_{Xt} \right) + \left(\frac{1}{\gamma} - 1 \right) [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \mathbf{M}_1]' (\boldsymbol{\Sigma}'_P)^{-1} \boldsymbol{\Lambda}_P \right] F_\lambda \\ & \left. + \frac{1}{2} [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \mathbf{M}_1]' [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_{\mathcal{I}}^{-1} \mathbf{M}_1] F_{\lambda\lambda}, \right. \end{aligned} \quad (\text{B.2.5})$$

with boundary condition $F(\lambda_{Xt}, T) = 1$. The solution to the differential equation is

$$F(\hat{\lambda}_{Xt}, t) = \exp \left\{ \left(\frac{1}{\gamma} - 1 \right) \left(A(t) + B(t) \hat{\lambda}_{Xt} + \frac{1}{2} C(t) \hat{\lambda}_{Xt}^2 \right) \right\} \quad (\text{B.2.6})$$

where $A(t)$, $B(t)$, and $C(t)$ are deterministic functions of time, with $A(T) = B(T) = C(T) = 0$.

To verify that the solution in (B.2.6) satisfies the partial differential equation in (B.2.5), it is convenient to split the risk premium vector $\boldsymbol{\Lambda}_P$ into two components:

$$\begin{aligned} \boldsymbol{\Lambda}_P &= \begin{bmatrix} \lambda_{Xt} \\ \lambda_{Xt} + \lambda_S + \sigma_X \sigma_s \rho_{XS} \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda_S + \sigma_X \sigma_s \rho_{XS} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \lambda_{Xt} \\ &= \mathbf{L}_0 + \mathbf{L}_1 \lambda_{Xt}. \end{aligned} \quad (\text{B.2.7})$$

Substituting (B.2.6) into (B.2.5) and simplifying leads to an equation that is quadratic in λ_{Xt} , i.e. it takes the form

$$0 = f(t) + g(t) \lambda_{Xt} + h(t) \lambda_{Xt}^2. \quad (\text{B.2.8})$$

This equation must hold for any value of λ_{Xt} . Setting the coefficients equal to zero

leads to a system of ordinary differential equations,

$$\begin{aligned} \frac{dC}{dt} = & -\frac{1}{\gamma} \mathbf{L}'_1 (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P)^{-1} \mathbf{L}_1 \\ & + 2 \left[\kappa_\lambda - \left(\frac{1}{\gamma} - 1 \right) \left(\boldsymbol{\Sigma}'_\lambda \boldsymbol{\Sigma}_P^{-1} \mathbf{L}_1 + \mathbf{M}'_1 (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}_I)^{-1} \mathbf{L}_1 \nu(t) \right) \right] C(t) \\ & - \left(\frac{1}{\gamma} - 1 \right) [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_I^{-1} \mathbf{M}_1]' [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_I^{-1} \mathbf{M}_1] C(t)^2 \end{aligned} \quad (\text{B.2.9a})$$

$$\begin{aligned} \frac{dB}{dt} = & -\frac{1}{\gamma} \mathbf{L}'_0 (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P)^{-1} \mathbf{L}_1 \\ & + \left[\kappa_\lambda - \left(\frac{1}{\gamma} - 1 \right) \left(\boldsymbol{\Sigma}'_\lambda \boldsymbol{\Sigma}_P^{-1} \mathbf{L}_1 + \mathbf{M}'_1 (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}_I)^{-1} \mathbf{L}_1 \nu(t) \right) \right] B(t) \\ & - \left[\frac{1}{2} \kappa_\lambda \sigma_X^2 + \left(\frac{1}{\gamma} - 1 \right) \left(\boldsymbol{\Sigma}'_\lambda \boldsymbol{\Sigma}_P^{-1} \mathbf{L}_0 + \mathbf{M}'_1 (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}_I)^{-1} \mathbf{L}_0 \nu(t) \right) \right] C(t) \\ & - \left(\frac{1}{\gamma} - 1 \right) [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_I^{-1} \mathbf{M}_1]' [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_I^{-1} \mathbf{M}_1] B(t) C(t) \end{aligned} \quad (\text{B.2.9b})$$

$$\begin{aligned} \frac{dA}{dt} = & -r_d - \frac{1}{\gamma} \mathbf{L}'_0 (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}'_P)^{-1} \mathbf{L}_0 \\ & - \left[\frac{1}{2} \kappa_\lambda \sigma_X^2 + \left(\frac{1}{\gamma} - 1 \right) \left(\boldsymbol{\Sigma}'_\lambda \boldsymbol{\Sigma}_P^{-1} \mathbf{L}_0 + \mathbf{M}'_1 (\boldsymbol{\Sigma}_P \boldsymbol{\Sigma}_I)^{-1} \mathbf{L}_0 \nu(t) \right) \right] B(t) \\ & - \frac{1}{2} \left(\frac{1}{\gamma} - 1 \right) [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_I^{-1} \mathbf{M}_1]' [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_I^{-1} \mathbf{M}_1] B(t)^2 \\ & - \frac{1}{2} [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_I^{-1} \mathbf{M}_1]' [\boldsymbol{\Sigma}_\lambda + \nu(t) \boldsymbol{\Sigma}_I^{-1} \mathbf{M}_1]. \end{aligned} \quad (\text{B.2.9c})$$

Together with (B.1.1), the equations in (B.2.9) describe a system of ordinary differential equations that can be solved sequentially. I solve the system numerically using the fourth-order Runge-Kutta method.