NBER WORKING PAPER SERIES

TECHNOLOGICAL REVOLUTIONS AND STOCK PRICES

Lubos Pastor
Pietro Veronesi

Working Paper 11876
http://www.nber.org/papers/w11876

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2005

Both authors are at the Graduate School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA. Email: lubos.pastor@ChicagoGSB.edu and pietro.veronesi@ChicagoGSB.edu.
We thank Malcolm Baker, Robert Barro, Efraim Benmelech, Markus Brunnermeier, Gene Fama, Bob Fogel, Boyan Jovanovic, John Heaton, Robert Novy-Marx, Rob Stambaugh, Dmitriy Stolyarov, and the audience participants at the Fall 2005 NBER Asset Pricing meeting, CERGE-EI (Prague), Dartmouth College, Ente Einaudi (Rome), Indiana University, London Business School, London School of Economics, New York University, Stockholm Institute for Financial Research, Stockholm School of Economics, and University of Chicago for helpful comments. We also thank Shastri Sandy for valuable research assistance.The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Technological Revolutions and Stock Prices
Lubos Pastor and Pietro Veronesi
NBER Working Paper No. 11876
December 2005
JEL No. G1

ABSTRACT

During technological revolutions, stock prices of innovative firms tend to exhibit high volatility and bubble-like patterns, which are often attributed to investor irrationality. We develop a general equilibrium model that rationalizes the observed price patterns. The high volatility results from high uncertainty about the average productivity of a new technology. Investors learn about this productivity before deciding whether to adopt the technology on a large scale. For technologies that are ultimately adopted, the nature of uncertainty changes from idiosyncratic to systematic as the adoption becomes more likely; as a result, stock prices fall after an initial run-up. This "bubble" in stock prices is observable ex post but unpredictable ex ante, and it is most pronounced for technologies characterized by high uncertainty and fast adoption. We examine stock prices in the early days of American railroads, and find evidence consistent with a large-scale adoption of the railroad technology by the late 1850s.

Lubos Pastor
Graduate School of Business
University of Chicago
5807 South Woodlawn Ave
Chicago, IL 60637
and NBER
lubos.pastor@chicagogsb.edu

Pietro Veronesi
Graduate School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
pietro.veronesi@chicagogsb.edu

"Technological revolutions and financial bubbles seem to go hand in hand." The Economist, Sep 21, 2000.

1. Introduction

Technological revolutions tend to be accompanied by bubble-like patterns in the stock prices of firms that employ the new technology. After an initial surge, stock prices of innovative firms usually fall in the presence of high volatility. Recent examples of such price patterns include the "Internet craze" of the late 1990s, the "biotech revolution" of the early 1980s, and the "tronics boom" of the early 1960s, as characterized by Malkiel (1999).[1] Other examples include the 1920s and the turn of the 20th century; in both periods, technological innovation spread rapidly while the stock market boomed and then faltered (e.g., Shiller, 2000).[2]

The bubble-like stock price behavior during technological revolutions is frequently attributed to market irrationality (e.g., Shiller, 2000, Perez, 2002). For example, one common argument is that investors repeatedly fail to realize that technological advances benefit mostly consumers rather than producers. In this paper, we propose an alternative explanation, without appealing to irrationality. We argue that new technologies are characterized by high uncertainty about their average future productivity, and that the time-varying nature of this uncertainty can produce the observed stock price patterns.

We build a general equilibrium model of a finite-horizon representative-agent economy with two sectors: the "new economy" and the "old economy." The old economy implements the existing technologies in large-scale production whose cumulative output determines the representative agent's terminal wealth. The new economy, which is created when a new technology is invented, implements the new technology in small-scale production that does not affect the agent's wealth. Under simple assumptions, it is optimal for a new technology to be initially employed on a small scale because its future productivity is uncertain. By observing the new economy, the representative agent learns about the average productivity of the new technology before deciding (as a utility-maximizing social planner) whether to adopt the technology on a large scale. We show that this irreversible adoption takes place if the agent learns that the new technology is sufficiently productive. We define a technological revolution as a period concluded by a large-scale adoption of a new technology.

[1] According to Malkiel (1999), "What electronics was to the 1960s, biotechnology became to the 1980s... Valuation levels of biotechnology stocks reached levels previously unknown to investors... From the mid-1980s to the late 1980s, most biotechnology stocks lost three-quarters of their market value."

[2] According to The Economist (2000), "Every previous technological revolution has created a speculative bubble... With each wave of technology, share prices soared and later fell... The inventions of the late 19th century drove p-e ratios to a peak in 1901, the year of the first transatlantic radio transmission. By 1920 shares prices had dropped by 70% in real terms. The roaring twenties were also seen as a "new era": share prices soared as electricity boosted efficiency and car ownership spread. After peaking in 1929, real share prices tumbled by 80% over the next three years."

The nature of the risk associated with new technologies changes over time. Initially, this risk is mostly idiosyncratic due to the small scale of production and a low probability of adoption. The risk remains largely idiosyncratic for those technologies that are never adopted on a large scale. For the technologies that are ultimately adopted, however, the risk gradually changes from idiosyncratic to systematic: As the probability of adoption increases, the new technology becomes more likely to affect the old economy and with it the representative agent's wealth, so the systematic risk in the economy increases.

This time-varying nature of risk has interesting implications for stock prices. Initially, while uncertainty about the average productivity of the new technology is idiosyncratic, it increases both the level and volatility of stock prices in the new economy.[3] Therefore, the new economy stocks initially command high valuation ratios (of market value to book value of equity) and high volatility. However, as the adoption probability increases, the resulting increase in systematic risk increases the discount rates and thus depresses stock prices. Stock prices fall in both the new and old economies, but especially in the new economy. In short, we argue that stock prices begin falling during technological revolutions when it becomes likely that the new technology will eventually be adopted on a large scale.

Stock prices are affected not only by news about discount rates but also by news about cash flows. The technologies that are ultimately adopted must turn out to be sufficiently productive before the adoption. This positive cash flow news pushes stock prices up, countervailing the effect of the higher discount rate. The cash flow effect tends to prevail initially, pushing the new economy stock prices up, but the discount rate effect prevails eventually, pushing the stock prices down. The resulting pattern in the new economy stock prices looks like a bubble although it is perfectly rational.

The bubble-like pattern in stock prices can be viewed as an outcome of a "hindsight bias." Researchers study technological revolutions with the ex post knowledge that the revolutions took place, but investors living through those periods did not know whether the new technologies would eventually be adopted on a large scale. The representative agent in our model never expects stock prices to fall; she always expects to earn positive stock returns commensurate to the stocks' riskiness, and she subsequently earns those fair returns, on average. However, in those periods that are recognized as technological revolutions ex post, the agent's realized returns tend to be initially positive due to good news about productivity and eventually negative due to unexpected increases in systematic risk.

[3] Uncertainty about average productivity increases market value because the latter is convex in average productivity, as explained in Pástor and Veronesi (2003, 2006).

In addition to the level of stock prices, the high stock volatility observed during technological revolutions can also be explained by uncertainty about new technologies. Due to this uncertainty, the new economy stocks are more volatile than the old economy stocks. Moreover, the new economy's volatility exhibits a U-shape pattern: it initially declines due to learning, but it ultimately increases when the volatility of the stochastic discount factor increases as a result of a higher probability of a large-scale adoption. The latter effect also causes the old economy's volatility to increase during technological revolutions, albeit more slowly. We also show that if agents have to choose which new technologies to implement in the new economy, they prefer technologies with higher uncertainty. In that sense, uncertainty about average productivity is a natural feature of new technologies.

To complement our theoretical analysis, we empirically examine the stock price behavior during the first major technological revolution in the United States since the opening of the U.S. stock market – the introduction of steam-powered railroads. We argue that in the 1830s and 40s, there was substantial uncertainty about whether the railroad technology would be adopted on a large scale. We analyze stock prices before the Civil War, and find that they fell before and during year 1857, with railroad stocks falling more than non-railroad stocks. We also find that railroad stock volatility and price-dividend ratios consistently exceeded their non-railroad counterparts, and that the volatility of all stocks rose in 1857. In the context of our model, this evidence is consistent with a large-scale adoption of the railroad technology around 1857, after railroads began expanding west of the Mississippi River.

Much of the literature on technological innovation analyzes issues different from those addressed here. Unlike Romer (1990), Aghion and Howitt (1992), and others, we take technological inventions to be exogenous. We do not examine the links between technological revolutions and human capital (e.g., Chari and Hopenhayn, 1991, Caselli, 1999, Manuelli, 2003). Although there is learning in our model, there is no learning-by-doing in the sense of Arrow (1962), Jovanovic and Nyarko (1996), Atkeson and Kehoe (2003), and others because learning here does not affect the technology's productivity. In Jovanovic (1982), firms learn about their costs; the efficient firms grow, the inefficient ones decline. Our model is similar in that we learn about the average productivity of a new technology; the productive technologies are adopted, the unproductive ones are not. We empirically examine the "railroad revolution" in the U.S., while other technological revolutions are examined by Jovanovic and Rousseau (2003, 2005), Mazzucato (2002), and others. Mokyr (1990) argues that technological progress is discontinuous, as assumed in our model, and that occasional seminal inventions ("macroinventions") are the key sources of economic growth.

A small but growing literature explores the links between technological innovation and stock prices (e.g., Jovanovic and MacDonald, 1994, and Laitner and Stolyarov, 2003, 2004a,b). According to Greenwood and Jovanovic (1999) and Hobijn and Jovanovic (2001), innovation causes the stock market to drop because the incumbent firms are unable or unwilling to implement the new technology. Similar initial stock market drops are obtained in the models of Laitner and Stolyarov (2003) and Manuelli (2003). In our model, the stock market value of the old economy also drops after the new technology is invented, mostly because of the costs and risks associated with a large-scale adoption of the new technology, but our focus is on the subsequent bubble-like stock price pattern in the new economy.

We focus not only on the level of stock prices, as the above papers do, but also on the stock price volatility. Mazzucato (2002) studies the early phases of the life-cycles of the automobile and PC industries in the U.S., and finds that in both industries, stock prices were the most volatile when technological change was the most radical. Agarwal et al. (2004) empirically examine a sample of brick-and-mortar firms that launched Web sites in the late 1990s, and find that initiating eCommerce increases the idiosyncratic return volatility of a firm's stock. To explain their results, both papers argue that taking up the new technology increases the uncertainty that firms face in the product market. Both the empirical evidence and the explanations provided in these papers are consistent with our model.

The paper is organized as follows. Section 2 presents the model. Section 3 solves for stock prices and analyzes their dynamics. Section 4 calibrates the model and uses simulations to investigate the model-implied paths of stock prices and volatilities during technological revolutions. Section 5 empirically examines the behavior of stock prices in 1830 through 1861 when the railroad technology spread in the United States. Section 6 concludes.

2. The Economy

We consider an economy with a finite horizon $[0, T]$. A representative agent has preferences defined by power utility over final wealth W_T, with risk aversion $\gamma > 1$:

$$u(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}. \tag{1}$$

At time $t = 0$, the agent is endowed with capital B_0. Subsequently, capital is invested in a linear technology producing output (net of depreciation) at the rate of

$$Y_t = \rho_t B_t.$$

Since there is no intermediate consumption, all output is reinvested, and capital follows

$$dB_t = Y_t dt = \rho_t B_t dt. \tag{2}$$

Productivity ρ_t follows a mean-reverting process whose mean is determined by the available technology. There are two technologies: "old" and "new." Initially, only the old technology is available, and the long-run mean of ρ_t is equal to $\overline{\rho}$. At time t^*, the new technology becomes available. If the representative agent adopts the new technology at time $t^{**} \geq t^*$, the long-run mean of ρ_t changes from $\overline{\rho}$ to $\overline{\rho} + \psi$. Thus, the dynamics of ρ_t are given by

$$d\rho_t = \phi \left(\overline{\rho} - \rho_t\right) dt + \sigma dZ_{0,t}, \qquad 0 < t < t^{**} \tag{3}$$

$$d\rho_t = \phi \left(\overline{\rho} + \psi - \rho_t\right) dt + \sigma dZ_{0,t}, \qquad t^{**} \leq t < T, \tag{4}$$

where ϕ is the speed of mean reversion, $\overline{\rho}$ is the mean productivity of the old technology, ψ is the "productivity gain" brought by the new technology, and σ^2 is the variance of productivity shocks, represented by the Brownian increments $dZ_{0,t}$. That is, we define the adoption of the new technology simply as a shift in the economy's average productivity.

The representative agent chooses whether and when to adopt the new technology to maximize utility in equation (1) under the market-clearing condition $W_T = B_T$. In equilibrium, the agent's final wealth must equal the amount of capital accumulated by time T.

Our key assumption is that the productivity gain ψ is unobservable. When the new technology appears at time t^*, ψ is drawn from a normal distribution with known variance:

$$\psi \sim N\left(0, \widehat{\sigma}^2_{t^*}\right). \tag{5}$$

All other parameters are known. The adoption of the new technology is irreversible; after the adoption, the agent cannot go back to the old technology. Finally, converting capital to the new technology is costly, incurring a proportional conversion cost $\kappa \geq 0$.

Proposition 1: It is never optimal to adopt the new technology immediately at time t^*.

Adopting the new technology is risky – it may increase or decrease average productivity, depending on the sign of ψ. Since the representative agent is risk averse and the prior in equation (5) is centered at zero, immediate adoption of the new technology is suboptimal.[4]

To formalize this intuition, define the value function at time t as

$$V\left(B_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t, t; T\right) = E_t\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right], \tag{6}$$

[4]If the prior is centered at $\widehat{\psi}_{t^*} \neq 0$, Proposition 1 is modified so that it is not optimal to adopt the new technology at time t^* unless $\widehat{\psi}_{t^*}$ is sufficiently high. See Proposition 2 for an analogous relation.

where ρ_t follows the process in equation (4) and the representative agent's beliefs at time t are given by $\psi \sim N\left(\widehat{\psi}_t, \widehat{\sigma}_t^2\right)$. A closed-form expression for V is provided in Lemma A1 in the Appendix. The Appendix also shows that

$$V\left(B_{t^*}(1-\kappa), \rho_{t^*}, 0, \widehat{\sigma}_{t^*}^2, t^*; T\right) < V\left(B_{t^*}, \rho_{t^*}, 0, 0, t^*; T\right),$$

where the left-hand (right-hand) side captures expected utility conditional on adopting (not adopting) the new technology at time t^*.[5] This result holds for any κ, including $\kappa = 0$, as it is driven purely by the increase in risk resulting from the adoption of the new technology.

2.1. Learning in the New Economy

Although adopting the new technology immediately is suboptimal, it might become optimal later if the agent learns that ψ is high. The agent can learn about ψ by "experimenting" with the new technology – i.e., by implementing it on a small scale. As shown in Section 2.3., it is optimal for the agent to begin experimenting at time t^*, immediately after the new technology becomes available. After time t^*, the economy consists of two sectors: the small-scale "new economy," which employs the new technology, and the large-scale "old economy," whose productivity follows equation (3). The capital B_t^N used in the new economy is infinitely smaller than B_t, so the agent's wealth W_T is affected by the new technology only if this technology is adopted on a large scale (i.e., by the old economy). Denoting the new economy's productivity by ρ_t^N, the processes of B_t^N and ρ_t^N for $t > t^*$ are given by

$$dB_t^N = \rho_t^N B_t^N dt \tag{7}$$

$$d\rho_t^N = \phi\left(\overline{\rho} + \psi - \rho_t^N\right) dt + \sigma_{N,0} dZ_{0,t} + \sigma_{N,1} dZ_{1,t}, \tag{8}$$

where $Z_{1,t}$ is a Brownian motion uncorrelated with $Z_{0,t}$. The representative agent learns about ψ by observing ρ_t^N and ρ_t. The following lemma characterizes the learning process:

Lemma 1: Suppose the prior distribution of ψ at time t^* is normal, $\psi \sim N(0, \widehat{\sigma}_{t^*}^2)$. Then the posterior distribution of ψ at time t, $t^* < t < t^{**}$, conditional on $\mathcal{F}_t = \left\{\left(\rho_\tau^N, \rho_\tau\right) : t^* \leq \tau \leq t\right\}$ is also normal, $\psi | \mathcal{F}_t \sim N(\widehat{\psi}_t, \widehat{\sigma}_t^2)$, where the posterior mean $\widehat{\psi}_t$ follows the process

$$d\widehat{\psi}_t = \widehat{\sigma}_t^2 \frac{\phi}{\sigma_{N,1}} d\widetilde{Z}_{1,t}, \tag{9}$$

and the posterior variance $\widehat{\sigma}_t^2$ is given by

$$\widehat{\sigma}_t^2 = \frac{1}{\left(\widehat{\sigma}_{t^*}\right)^{-2} + \left(\frac{\phi}{\sigma_{N,1}}\right)^2 (t - t^*)}. \tag{10}$$

[5] On the right hand side, V is evaluated at $\widehat{\psi}_{t^*} = \widehat{\sigma}_{t^*}^2 = 0$. If the agent decides not to adopt the new technology, ρ_t follows the process in equation (3), which is equivalent to equation (4) when $\psi = 0$.

Note that the posterior variance $\widehat{\sigma}_t^2$ declines deterministically over time due to learning. Also due to learning, the shocks perceived by the representative agent are given by the orthogonalized Brownian motions $\left(\widetilde{Z}_{0,t}, \widetilde{Z}_{1,t}\right)$ capturing the agent's expectation errors (see the Appendix). The productivity processes can then be rewritten as

$$
\begin{aligned}
d\rho_t &= \phi\left(\overline{\rho} - \rho_t\right) dt + \sigma d\widetilde{Z}_{0,t} && (11) \\
d\rho_t^N &= \phi\left(\overline{\rho} + \widehat{\psi}_t - \rho_t^N\right) dt + \sigma_{N,0} d\widetilde{Z}_{0,t} + \sigma_{N,1} d\widetilde{Z}_{1,t}. && (12)
\end{aligned}
$$

2.2. Technological Revolution

We define a technological revolution as the period $[t^*, t^{**}]$ concluded by a large-scale adoption of a new technology. We treat the invention of the new technology as given, and study the conditions under which the invention leads to a technological revolution.

When the new technology becomes available at time t^*, the representative agent acquires a real option to adopt the technology anytime before time T. The agent begins learning about the technology's productivity gain in the new economy, and solves for the optimal time t^{**} to adopt the technology in the old economy. (Such an adoption may or may not occur.) We solve for the optimal t^{**} numerically in Section 4.2. Here, we focus on a simpler problem in which, at a given time t^{**}, the agent decides whether or not to adopt the new technology. This simplification leads to closed-form solutions for stock prices, and thus improves our understanding of the price dynamics during technological revolutions. Our numerical results in Section 4.2. show that the price patterns obtained when t^{**} is endogenously chosen are very similar to those obtained with an exogenous t^{**}.

Proposition 2: The new technology is adopted at time t^{**} if and only if

$$
\widehat{\psi}_{t^{**}} \geq \underline{\psi} = -\frac{\log\left(1-\kappa\right)}{A_2\left(\tau^{**}\right)} + \frac{1}{2}\left(\gamma - 1\right) A_2\left(\tau^{**}\right) \widehat{\sigma}_{t^{**}}^2 , \qquad (13)
$$

where $\tau^{**} = T - t^{**}$, $A_2(\tau) = \tau - (1 - \exp(-\phi\tau))/\phi > 0$, and $\widehat{\sigma}_t$ is defined in Lemma 1.

The new technology is adopted if the expected productivity gain $\widehat{\psi}_{t^{**}}$ is sufficiently large. The threshold $\underline{\psi} > 0$ increases in the conversion cost κ and uncertainty $\widehat{\sigma}_{t^{**}}$, which is intuitive. Using our closed-form expression for the value function in equation (6), equation (13) follows from the optimality condition

$$
V\left(B_{t^{**}}\left(1-\kappa\right), \rho_{t^{**}}, \widehat{\psi}_{t^{**}}, \widehat{\sigma}_{t^{**}}^2, t^{**}; T\right) \geq V\left(B_{t^{**}}, \rho_{t^{**}}, 0, 0, t^{**}; T\right). \qquad (14)
$$

Note that the agent makes the adoption decision without knowing for sure whether the new technology increases productivity. Regardless of the outcome of the adoption decision, the true value of ψ remains unknown and learning about ψ continues after time t^{**}.

2.3. Optimal Experimentation under Uncertainty

We now show that the agent sets up the new economy and begins learning about the new technology immediately after this technology becomes available at time t^*.

Proposition 3: It is optimal to begin experimenting with the new technology at time t^*.

To prove the proposition formally, define the value function at time t, $t^* \leq t < t^{**}$, as

$$\mathcal{V}\left(B_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2, t; T\right) = E_t \left\{ \max_{\{\text{yes, no}\}} E_{t^{**}} \left[\frac{W_T^{1-\gamma}}{1-\gamma} \right] \right\}, \quad (15)$$

where the maximization involves choosing whether or not to adopt the new technology at time t^{**}, following Proposition 2. The Appendix provides an expression for $\mathcal{V}$ (Lemma A3), along with a proof that expected utility is higher when experimentation takes place:

$$\mathcal{V}\left(B_{t^*}, \rho_{t^*}, 0, \widehat{\sigma}_{t^*}^2, t; T\right) > V\left(B_{t^*}, \rho_{t^*}, 0, 0, t; T\right). \quad (16)$$

The intuition behind Proposition 3 is simple. Experimenting allows the agent to learn about the productivity gain ψ. If this learning leads the agent to believe at time t^{**} that ψ is sufficiently high, then it becomes optimal to adopt the new technology (Proposition 2). Otherwise, the status quo will prevail. Since experimenting is costless and there is no downside to it, it gives the agent a valuable option for free.[6]

Since option value generally increases with uncertainty, high uncertainty $\widehat{\sigma}_{t^*}$ makes a new technology desirable for experimentation.[7] If it were costly to experiment with new technologies, or if the agent had to choose from a subset of technologies at time t^*, then the technologies with the highest $\widehat{\sigma}_{t^*}$ would be selected for experimentation, ceteris paribus. Uncertainty about productivity gains is thus a natural feature of innovative technologies.

The sequence of events in the model is summarized in Figure 1. We assume that if a new technology is not adopted at time t^{**}, it continues to operate on a small scale until time T.

[6] The problem we solve resembles the problem of making an irreversible marriage decision. It is generally suboptimal to marry a new acquaintance immediately because of substantial uncertainty regarding the quality of the personality match (cf. Proposition 1). Instead, it seems advisable to first develop the relationship on a small scale, by dating without any commitment (cf. Proposition 3), and then to marry if we learn that the relationship is likely to work in the long run (cf. Proposition 2).

[7] We find numerically that the value function $\mathcal{V}$ is increasing in $\widehat{\sigma}_{t^*}$ ($\partial\mathcal{V}/\partial\widehat{\sigma}_{t^*} > 0$) for any reasonable parameter values. In fact, we have not found any parameter values for which $\partial\mathcal{V}/\partial\widehat{\sigma}_{t^*} > 0$ is violated. While a general proof that $\partial\mathcal{V}/\partial\widehat{\sigma}_{t^*} > 0$ seems infeasible, we have some local analytical results. Proposition 3 shows that $\mathcal{V}$ is increasing in $\widehat{\sigma}_{t^*}$ as $\widehat{\sigma}_{t^*} \to 0$, and for $\kappa = 0$, we can also prove that $\partial\mathcal{V}/\partial\widehat{\sigma}_{t^*} > 0$ as $\widehat{\sigma}_{t^*} \to \infty$. Given $\partial\mathcal{V}/\partial\widehat{\sigma}_{t^*} > 0$, if we added an assumption that experimenting with new technologies is costly, Proposition 3 would be modified so that it is optimal to begin experimenting at time t^* unless $\widehat{\sigma}_{t^*}$ is too low.

Our history is full of examples of “failed” technologies that have not been adopted on a large scale but still survive on a small scale (e.g., direct-current electric motors, airships, etc.)

3. Stock Prices

The stocks of the old and new economies are the contingent claims paying liquidating dividends B_T and B_T^N, respectively, at time T. There is also a riskless bond in zero net supply, whose yield we normalize to zero, for simplicity. Since the two shocks in the model ($\widetilde{Z}_0$ and $\widetilde{Z}_1$) are spanned by the two stocks, markets are complete. Standard arguments then imply that the state price density is uniquely given by

$$\pi_t = \frac{1}{\lambda} E_t \left[W_T^{-\gamma} \right], \tag{17}$$

where λ is the Lagrange multiplier from the utility maximization problem of the representative agent. The market values (shadow prices) of the old and new economy stocks, denoted by M_t and M_t^N, respectively, are given by the standard pricing formulas

$$M_t = E_t \left[\frac{\pi_T B_T}{\pi_t} \right] \qquad \text{and} \qquad M_t^N = E_t \left[\frac{\pi_T B_T^N}{\pi_t} \right]. \tag{18}$$

To normalize the market values, we form “market-to-book” (M/B) ratios M_t/B_t and M_t^N/B_t^N. It seems reasonable to interpret capital as the book value of equity, and this interpretation is exact for B_t and B_t^N in equations (2) and (7) if we also interpret output and productivity as earnings and profitability, respectively (Pástor and Veronesi, 2003).

Let p_t denote the probability at time t, $t^* \leq t < t^{**}$, that the new technology will be adopted at time t^{**}. Lemma A2 in the Appendix shows that

$$p_t = 1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2 \right), \tag{19}$$

where $\mathcal{N}(\cdot; a, s^2)$ denotes the cumulative density function of the normal distribution with mean a and variance s^2, and $\widehat{\sigma}_t^2$ is given in Lemma 1.

Proposition 4: For any $t \in [t^*, t^{**})$, the state price density is given by

$$\pi_t = \lambda^{-1} B_t^{-\gamma} \left\{ (1 - p_t)\, \widetilde{G}_t^{no} + p_t \widetilde{G}_t^{yes} \right\}, \tag{20}$$

where

$$\widetilde{G}_t^{no} = E_t \left[\left(\frac{B_T}{B_t} \right)^{-\gamma} | \widehat{\psi}_{t^{**}} < \underline{\psi} \right] = e^{\overline{A}_0(\tau) - \gamma A_1(\tau) \rho_t} \tag{21}$$

$$\widetilde{G}_t^{yes} = E_t \left[\left(\frac{B_T}{B_t} \right)^{-\gamma} | \widehat{\psi}_{t^{**}} \geq \underline{\psi} \right], \tag{22}$$

and where $\tau = T - t$, $A_1(\tau) = (1 - e^{-\phi\tau})/\phi$, and $\overline{A}_0(\tau)$ and $\widetilde{G}_t^{yes}$ are in the Appendix.

Intuitively, π_t is a probability-weighted average of the expectations of marginal utility of wealth conditional on whether or not the new technology is adopted at time t^{**}. (Recall from Proposition 2 that the adoption takes place if $\widehat{\psi}_{t^{**}} \geq \underline{\psi}$, which occurs with probability p_t.) Computing $\widetilde{G}_t^{yes}$ is more complicated than computing $\widetilde{G}_t^{no}$ because the adoption of the new technology changes the dynamics of ρ_t from (3) to (4), which makes B_T depend on $\widehat{\psi}_{t^{**}}$.

Corollary 1. For any $t \in [t^*, t^{**})$, the dynamics of π_t are given by

$$\frac{d\pi_t}{\pi_t} = -\sigma_{\pi,t}^0 d\widetilde{Z}_{0,t} - \sigma_{\pi,t}^1 d\widetilde{Z}_{1,t} = -\gamma A_1(\tau)\sigma d\widetilde{Z}_{0,t} - S_{\pi,t}\widehat{\sigma}_t^2 \frac{\phi}{\sigma_{N,1}} d\widetilde{Z}_{1,t}, \tag{23}$$

where $S_{\pi,t}$ is given in the Appendix.

This corollary illustrates the time-varying nature of risk during technological revolutions. When a new technology arrives at time t^*, the adoption probability p_{t^*} is generally small, which makes S_{π,t^*} small as well ($p_t = 0$ implies $S_{\pi,t} = 0$). The volatility of the stochastic discount factor in equation (23) then depends only slightly on $\widehat{\sigma}_t^2$, making uncertainty about ψ mostly idiosyncratic. During a technological revolution, the adoption probability increases, which makes $S_{\pi,t}$ larger.[8] As a result, the volatility of the stochastic discount factor becomes more closely tied to $\widehat{\sigma}_t^2$, making uncertainty about ψ increasingly systematic.

Proposition 5: For any $t \in [t^*, t^{**})$, the market-to-book ratios are given by

$$\frac{M_t}{B_t} = \frac{(1-p_t)G_t^{no} + p_t G_t^{yes}}{(1-p_t)\,\widetilde{G}_t^{no} + p_t \widetilde{G}_t^{yes}} \tag{24}$$

$$\frac{M_t^N}{B_t^N} = \frac{(1-p_t)K_t^{no} + p_t K_t^{yes}}{(1-p_t)\,\widetilde{G}_t^{no} + p_t \widetilde{G}_t^{yes}}, \tag{25}$$

where $\widetilde{G}_t^{no}$ and $\widetilde{G}_t^{yes}$ are given in Proposition 4, and

$$G_t^{no} = E_t\left[\left(\frac{B_T}{B_t}\right)^{1-\gamma} |\widehat{\psi}_{t^{**}} < \underline{\psi}\right]; \qquad G_t^{yes} = E_t\left[\left(\frac{B_T}{B_t}\right)^{1-\gamma} |\widehat{\psi}_{t^{**}} \geq \underline{\psi}\right] \tag{26}$$

$$K_t^{no} = E_t\left[\left(\frac{B_T}{B_t}\right)^{-\gamma} \frac{B_T^N}{B_t^N}|\widehat{\psi}_{t^{**}} < \underline{\psi}\right]; \qquad K_t^{yes} = E_t\left[\left(\frac{B_T}{B_t}\right)^{-\gamma} \frac{B_T^N}{B_t^N}|\widehat{\psi}_{t^{**}} \geq \underline{\psi}\right], \tag{27}$$

are given explicitly in the Appendix.

[8]The dependence of $S_{\pi,t}$ on p_t is difficult to characterize explicitly because both variables depend on $\widehat{\psi}$. Although the dependence need not be monotonic, $S_{\pi,t}$ generally increases as p_t increases. At time t^*, we have $p_{t^*} \approx 0$ and $S_{\pi,t^*} \approx 0$. In a technological revolution, p_t rises to $p_{t^{**}} = 1$, at which point $S_{\pi,t^{**}} = \gamma A_2(\tau^{**}) > 0$. That is, as p_t increases, $S_{\pi,t}$ increases from approximately zero to a positive number.

In the special case $p_t = 0$, the market-to-book ratio of the new economy simplifies into

$$\frac{M_t^N}{B_t^N} = e^{C_0(\tau)+A_1(\tau)\rho_t^N+A_2(\tau)\widehat{\psi}_t+\frac{1}{2}A_2(\tau)^2\widehat{\sigma}_t^2}, \tag{28}$$

where $A_1(\tau)$ is defined in Proposition 4, $A_2(\tau)$ in Proposition 2, and $C_0(\tau)$ is in the Appendix. Note that M^N/B^N increases when uncertainty about ψ, $\widehat{\sigma}_t^2$, increases. This relation, first pointed out by Pástor and Veronesi (2003) in a simpler framework, is due to the idiosyncratic nature of uncertainty. When $p_t = 0$, the state price density does not depend on uncertainty about ψ, but when $p_t > 0$, it does.[9] When p_t is sufficiently large, uncertainty is mostly systematic, and the associated risk reverses the positive relation between M^N/B^N and $\widehat{\sigma}_t^2$.[10]

Corollary 2: For any $t \in [t^*, t^{**})$, the stock return processes are given by

$$\frac{dM_t}{M_t} = \mu_{M,t}dt + \sigma_{M,t}^0 d\widetilde{Z}_t^0 + \sigma_{M,t}^1 d\widetilde{Z}_t^1 \quad \text{and} \quad \frac{dM_t^N}{M_t^N} = \mu_{M,t}^N dt + \sigma_{M,t}^{N,0} d\widetilde{Z}_t^0 + \sigma_{M,t}^{N,1} d\widetilde{Z}_t^1,$$

where expected returns are equal to the return covariances with $d\pi_t/\pi_t$,

$$\mu_{M,t} = -\sigma_{M,t}^0\sigma_{\pi,t}^0 - \sigma_{M,t}^1\sigma_{\pi,t}^1 \tag{29}$$

$$\mu_{M,t}^N = -\sigma_{M,t}^{N,0}\sigma_{\pi,t}^0 - \sigma_{M,t}^{N,1}\sigma_{\pi,t}^1, \tag{30}$$

and the components of the return volatilities are

$$\sigma_{M,t}^0 = A_1(\tau)\,\sigma; \qquad \sigma_{M,t}^1 = (S_{M,t} + S_{\pi,t})\,\widehat{\sigma}_t^2\frac{\phi}{\sigma_{N,1}} \tag{31}$$

$$\sigma_{M,t}^{N,0} = A_1(\tau)\,\sigma_{N,0}; \qquad \sigma_{M,t}^{N,1} = A_1(\tau)\,\sigma_{N,1} + \left(S_{M,t}^N + S_{\pi,t}\right)\widehat{\sigma}_t^2\frac{\phi}{\sigma_{N,1}}, \tag{32}$$

with $S_{M,t}$ and $S_{M,t}^N$ given in the Appendix.

Note that the return volatilities in both economies increase with uncertainty $\widehat{\sigma}_t^2$.

3.1. The Dynamics of Prices during a Technological Revolution

In a technological revolution, the adoption probability p_t rises from a small value at time t^* to the value of one at time t^{**}. The effect of p_t on stock prices is analyzed next.

Proposition 6: The new (old) economy's market-to-book ratio is increasing in p_t if and only if $h_{new} > 0$ ($h_{old} > 0$), where the functions h_{new} and h_{old} are given in the Appendix.

[9] When $p_t = 0$, the state price density in equation (20) simplifies into $\pi_t = \lambda^{-1}B_t^{-\gamma}\exp\{\overline{A}_0(\tau) - \gamma A_1(\tau)\rho_t\}$.

[10] These results hold also in a more general model in which ψ is not constant (an assumption we make for simplicity) but rather decays gradually toward zero. In this alternative specification, the level of M^N/B^N is lower. However, the bubble-like dynamics of stock prices, which we document below, are unaffected unless the rate of decay in ψ is too large (in which case uncertainty about ψ becomes unimportant).

To illuminate the conditions derived in Proposition 6, Figure 2 plots h_{new} and h_{old} as functions of $\widehat{\psi}_t$ and $t^{**} - t$, for the parameter values used in our subsequent calibration. Panel A shows that the condition $h_{new} > 0$ is satisfied when $\widehat{\psi}_t$ is close to its initial value of zero ($\widehat{\psi}_{t^*} = 0$), but the condition becomes violated as $\widehat{\psi}_t$ increases towards the threshold $\underline{\psi}$ (marked by a dotted vertical line). In addition, h_{new} turns negative as $t^{**} - t$ falls, holding $\widehat{\psi}_t$ constant. In other words, $h_{new} > 0$ holds initially, shortly after time t^*, but it becomes violated as time t^{**} approaches and the adoption at time t^{**} becomes more likely. Proposition 6 then implies that the new economy's M/B is initially increasing but ultimately decreasing in p_t in the course of a technological revolution.

Panel B of Figure 2 shows that the condition $h_{old} > 0$ is never satisfied for the given parameter values, so the old economy's M/B is always decreasing in p_t. Note that h_{old} increases when $\widehat{\psi}_t$ increases because adopting a new technology is more valuable when the technology is more productive. Additional analysis shows that increases in κ or $\widehat{\sigma}_t$ lead to decreases in h_{old} because adoption that involves higher conversion costs or a higher discount rate is less desirable. The condition $h_{old} > 0$ can be satisfied if κ and $\widehat{\sigma}_t$ are sufficiently small and $\widehat{\psi}_t$ is sufficiently large, but for most reasonable parameter values, $h_{old} < 0$.

While analyzing M/B as a function of p_t seems informative, p_t itself is driven primarily by $\widehat{\psi}_t$. Stock prices depend on $\widehat{\psi}_t$ through two channels working in opposite directions. On one hand, an increase in $\widehat{\psi}_t$ is good news for prices because it increases expected cash flows in both economies ($E_t[B_T]$ and $E_t\left[B_T^N\right]$). This *cash flow effect* is stronger for the new economy whose perceived productivity is immediately affected; the old economy's productivity is not affected by ψ until time t^{**}, if at all. On the other hand, an increase in $\widehat{\psi}_t$ is bad news for prices because the higher adoption probability makes the risk embedded in the new technology increasingly systematic, thereby raising the discount rate. This *discount rate effect* is also stronger for the new economy because the stochastic discount factor covaries more with ρ_t^N than with ρ_t (since both $d\pi_t/\pi_t$ and ρ_t^N depend on $\widetilde{Z}_1$, but ρ_t does not). Moreover, the discount rate effect has a growing impact on the new economy's M/B because the dependence of $d\pi_t/\pi_t$ on $\widetilde{Z}_1$ increases as p_t increases (equation (23)). For the old economy, the discount rate effect generally outweighs the cash flow effect from the very beginning, leading to a gradual price decline during a revolution. For the new economy, the cash flow effect tends to dominate at first, but the discount rate effect dominates in the end, producing a bubble-like pattern in the new economy stock prices.

Although characterizing the dependence of M^N/B^N on $\widehat{\psi}_t$ seems intractable in general, its key features can be established locally at times t^* and t^{**}. We show below that M^N/B^N

is increasing (decreasing) in $\widehat{\psi}$ around time t^* (t^{**}), under certain assumptions.

Proposition 7: For any $t \geq t^*$ there exists $\bar{p} > 0$ such that if $p_t < \bar{p}$ then $\frac{\partial\left(M_t^N/B_t^N\right)}{\partial\widehat{\psi}_t} > 0$.

In words, if the probability of adoption p_t is sufficiently small, then M^N/B^N is increasing in $\widehat{\psi}$. When p_t is close to zero, so is its sensitivity to changes in $\widehat{\psi}_t$; thus an increase in $\widehat{\psi}_t$ does not produce a large discount rate effect. The cash flow effect is large, though, because M^N/B^N in equation (28) is strongly increasing in $\widehat{\psi}$. Proposition 7 follows.

When a new technology arrives at time t^*, its probability of eventual adoption is typically small because only a small fraction of new technologies are adopted by the whole economy. Proposition 7 then implies that, for most new technologies, the cash flow effect initially prevails over the discount rate effect and M^N/B^N is increasing in $\widehat{\psi}$ shortly after time t^*.

We also have some local results at time t^{**}. Below, we compare the M/B ratio of the new economy under two scenarios: $\widehat{\psi}_{t^{**}} = \underline{\psi} \pm \varepsilon$, where $\varepsilon > 0$ is small.

Corollary 3:
(a) If $\widehat{\psi}_{t^{**}} = \underline{\psi} + \varepsilon$, then the new technology is adopted at time t^{**}, and

$$\frac{M_{t^{**}}^N}{B_{t^{**}}^N} = e^{\overline{C}_0(\tau^{**}) + A_1(\tau^{**})\rho_{t^{**}}^N + A_2(\tau^{**})\widehat{\psi}_{t^{**}} + \frac{1}{2}A_2(\tau^{**})^2(1-2\gamma)\widehat{\sigma}_{t^{**}}^2}. \tag{33}$$

(b) If $\widehat{\psi}_{t^{**}} = \underline{\psi} - \varepsilon$, then the new technology is not adopted at time t^{**}, and

$$\frac{M_{t^{**}}^N}{B_{t^{**}}^N} = e^{\overline{C}_0(\tau^{**}) + A_1(\tau^{**})\rho_{t^{**}}^N + A_2(\tau^{**})\widehat{\psi}_{t^{**}} + \frac{1}{2}A_2(\tau^{**})^2\widehat{\sigma}_{t^{**}}^2}. \tag{34}$$

The M/B of the new economy is clearly lower when the technological revolution takes place. The reason is the uncertainty term $\widehat{\sigma}_t^2$, whose coefficient is negative in part (a) and positive in part (b). In part (a), $\widehat{\sigma}_t^2$ is systematic (it affects π_t), whereas in part (b), it is idiosyncratic (it does not affect π_t). Since $\widehat{\psi}_t$ is essentially the same in both scenarios, the difference between M/B in parts (a) and (b) is due to the discount rate effect.

Close to the adoption time t^{**}, the discount rate effect is generally stronger than the cash flow effect. For the cash flow effect to prevail in the knife-edge case discussed above, $\widehat{\psi}_t$ would have to increase by at least $\gamma A_2(\tau^{**})\widehat{\sigma}_{t^{**}}^2$ to offset the higher systematic risk resulting from the adoption. Such an increase in $\widehat{\psi}_t$ seems implausibly large, given the parameters used in our calibration. Since the discount rate effect dominates, M^N/B^N decreases in $\widehat{\psi}_t$.

In summary, the cash flow effect usually dominates close to time t^*, leading to an initial positive relation between M^N/B^N and $\widehat{\psi}_t$, but the discount rate effect usually dominates close to time t^{**}, leading to an eventual negative relation. During a technological revolution, $\widehat{\psi}_t$ generally increases, leading to a bubble-like pattern in M^N/B^N.

3.2. Discussion

Corollary 3 shows that the adoption reduces the new economy's M/B, holding $\widehat{\psi}_t$ constant. Intuitively, the adoption of the new technology by the old economy does not bring any benefit to the new economy, which already uses the new technology. On the contrary, the adoption (or even an increasing probability thereof) increases systematic risk and thus reduces the new economy's market value. It appears that the adoption is not favored by the new economy shareholders. However, in the model, there is only one shareholder, the representative agent, who employs infinitely more capital in the old economy than in the new economy. This agent wants the adoption to take place because the utility gain from making the old economy more productive outweighs the (negligible) loss of market value in the new economy.

Analogous to Corollary 3, we can show that the old economy's market value also decreases at time t^{**} if the adoption takes place when $\widehat{\psi}_{t^{**}}$ is close to $\underline{\psi}$. Interestingly, the representative agent chooses to adopt the new technology even if doing so reduces the market value of her stocks. There is a difference between maximizing utility and maximizing market value. The adoption occurs only if it increases the agent's expected utility. This adoption changes the economic environment by installing (what the agent perceives to be) a more productive technology and by increasing expected stock returns. In this new environment, stock prices are lower (due to higher discount rates) but expected utility is higher (due to higher expected wealth). Expected utility and stock prices need not move in the same direction because stock prices are related to the agent's marginal utility rather than to the level of utility.

We solve the social planner's problem in which a utility-maximizing representative agent owns all output by holding the stocks of the old and new economies. When a new technology is invented, it becomes property of the social planner. The social planner finds it optimal to set up a (small-scale) new economy to learn about the new technology before deciding whether to adopt this technology in the (large-scale) old economy. Upon adoption, there is no transfer from the old economy to the new economy because the new economy does not own the new technology (the social planner does). As an example of a new economy firm, Amazon was an early user of the Internet but it did not own the Internet technology.

As an alternative to the social planner's problem, one can analyze a competitive economy in which firms independently decide whether and when to adopt the new technology while maximizing their own market values. Although the decentralized problem does not seem to have a tractable solution for stock prices, not even with exogenous t^{**}, we believe that it would lead to similar price dynamics as the (tractable) social planner's problem. Suppose that a continuum of firms facing different conversion costs observe signals about ψ. As $\widehat{\psi}_t$ increases during a technological revolution, the proportion of firms that adopt the new technology also increases. This proportion might play the same role as the adoption probability in our model: As the proportion increases from (close to) zero to one, the volatility of the stochastic discount factor also increases, making the uncertainty about ψ increasingly systematic. The decentralized model can be analyzed in future work.

4. Empirical Implications

The purpose of this section is to analyze the model-implied paths of stock prices and volatilities during technological revolutions. We simulate 50,000 samples of shocks in our economy and compute the paths of the M/B ratios and volatilities in each simulated sample. We split the 50,000 samples into two groups, depending on whether or not the new technology was adopted at time t^{**}, and plot the average paths of prices and volatilities across all samples within each group. Our objective is to understand how these paths differ depending on whether or not the new technology leads to a technological revolution.

Table 1 shows the parameters used in our simulations. For the productivity processes, we choose parameters close to those estimated by Pástor and Veronesi (2006) for the dynamics of profitability. We equate productivity with profitability because all output in our model represents firm profits. The parameter values for the conversion cost, time horizon, risk aversion, and prior beliefs about ψ are varied later in our sensitivity analysis.

Figure 3 plots the average paths of $\widehat{\psi}_t$, p_t, and $\sigma_\pi \equiv \mathrm{Std}(d\pi_t/\pi_t)$. Panel A shows that the average drift in $\widehat{\psi}_t$ during technological revolutions is positive, due to conditioning on the ex post event that $\widehat{\psi}_{t^{**}} \geq \underline{\psi}$ (without such conditioning, $\widehat{\psi}_t$ is a martingale; see equation (9)).[11] Analogously, conditional on $\widehat{\psi}_{t^{**}} < \underline{\psi}$, $\widehat{\psi}_t$ in Panel B (no revolution) drifts downward. The drift is less pronounced in Panel B than in Panel A because $\widehat{\psi}_{t^*} = 0$ and $\underline{\psi} > 0$. The average probability of adoption, p_t, drifts up in Panel C (revolution) and down in Panel D (no revolution), as expected. The volatility of the stochastic discount factor, σ_π, is almost flat while p_t is low, but it increases as p_t increases (Panel E).

[11] Brown, Goetzmann and Ross (1995) provide a mathematical proof of a related statement in their analysis of stock returns conditional on the stock's survival through the end of the sample.

Figure 4 plots the average paths of M/B and volatility for the new economy (solid line) and the old economy (dashed line). The panels on the left are based on the samples in which $p_{t^{**}} = 1$ (revolution); the panels on the right condition on $p_{t^{**}} = 0$ (no revolution).[12] The dotted vertical lines mark the time when the new technology arrives, $t^* = 1$, and the time at which the agent decides whether to adopt the technology, $t^{**} = 9$.

Panel A of Figure 4 plots the average paths of M/B across all technological revolutions. The new economy's M/B exhibits a bubble-like pattern of an initial increase followed by a decrease, as predicted in Section 3.1. Since we are conditioning on the adoption of the new technology at time t^{**}, $\widehat{\psi}_t$ has increased between t^* and t^{**} (Figure 3). This increase in $\widehat{\psi}_t$ has two countervailing effects on prices. First, it increases expected future cash flow from the new technology, pushing M/B up. Second, it increases the adoption probability, which makes the risks associated with the new technology ever more systematic (affecting W_T), which then increases the discount rate applied to future cash flow, pushing M/B down. For the new economy, the cash flow effect is stronger at first, but the discount rate effect prevails in the end, producing an apparent bubble. For the old economy, the cash flow effect is weaker (i.e., $E_t\left[B_T\right]$ increases by less than $E_t\left[B_T^N\right]$) because the old economy's productivity is not affected by ψ until time t^{**}. As a result, the discount rate effect outweighs the cash flow effect from the outset, leading to a slow price decline in the old economy's M/B.

Different technological revolutions produce different paths in M/B, depending on the path of realized productivity. These individual paths look mostly like bubbles that peak at different times, and they are far less smooth than the average path plotted in Panel A of Figure 4. This average path shows that apparent bubbles are not merely possible in a rational world; they should in fact be expected during most technological revolutions.

Panel B of Figure 4 plots the average paths of M/B across all samples in which $p_{t^{**}} = 0$ (no revolution). In these samples, $\widehat{\psi}_t$ declines slightly between t^* and t^{**}, nudging the M/Bs down as well. The decline is larger in the new economy, for two reasons. One, the new economy's M/B is more sensitive to $\widehat{\psi}_t$, as discussed earlier. Two, uncertainty about ψ gradually declines due to learning, which reduces M/B for the new economy but not for the old economy (see equation (28)). Thanks in part to this uncertainty, the level of M/B is higher in the new economy than in the old economy, in both Panels A and B. Higher productivity is another reason why the new economy's M/B is higher in Panel A, even after time t^{**}. Although the adoption makes the long-run means of productivity equal in both

[12] The fraction of the simulated samples in which $p_{t^{**}} = 1$ is approximately equal to the ex ante probability of adoption implied by our parameter choices, $p_{t^*} = 7.56\%$, as expected.

economies, the productivity at time t^{**} is higher in the new economy ($\rho^N_{t^{**}}$ is likely to be high to make $\widehat{\psi}_{t^{**}} > \underline{\psi}$), lifting the M/B of the new economy above that of the old economy.

Panel C of Figure 4 plots the average paths of stock return volatility across all technological revolutions. Volatility is higher in the new economy than in the old economy, partly due to higher volatility of the fundamentals, but mostly due to uncertainty about ψ. To understand the U-shape in the new economy's volatility, recall that shocks to $\widehat{\psi}_t$ affect stock prices via the discount rate and cash flow effects, which work in opposite directions. Around time t^* (t^{**}), the cash flow (discount rate) effect dominates, so the two effects do not offset each other much and the volatility is high. The volatility is lowest when the two effects cancel each other, which happens at some point between times t^* and t^{**}; hence the U-shape. For the old economy, the discount rate effect dominates from the outset. As a result, the old economy's volatility slowly increases as the rising adoption probability makes the stochastic discount factor more volatile. The spike in volatility at time t^{**} is caused by high price variation in those simulated paths where $\widehat{\psi}_{t^{**}}$ is close to the adoption threshold $\underline{\psi}$. If $\widehat{\psi}_t$ is close to $\underline{\psi}$ as $t \to t^{**}$, then p_t swings between values close to zero and one, making returns highly volatile (Corollary 3). We show later that the volatility spike disappears (but all other effects remain) when t^{**} is chosen optimally instead of being fixed exogenously.

Panel D of Figure 4 plots the average return volatility across all no-revolution samples. In these samples, the adoption probability is mostly close to zero, so the discount rate effect is weak. Therefore, the cash flow shocks to stock prices are not offset much by the discount rate shocks, and the new economy's volatility is larger than in Panel C. The volatility increases slightly over time because the adoption probability falls from 7.56% at time t^* to zero at time t^{**}, making the discount rate effect progressively weaker, on average.

Note that our calibration produces plausible values for the level and volatility of stock prices. The new (old) economy's M/B in Figure 4 ranges from 4.1 to 8.4 (1.9 to 3.2), while the new (old) economy's volatility ranges from 22% to 41% (20% to 31%) per year. For comparison with a recent technology boom, when the technology-loaded Nasdaq index peaked in March 2000, its M/B stood at 8.55 and the standard deviation of its daily returns in March 2000 was 41% (Pástor and Veronesi, 2006). At the same time, the M/B ratio of the NYSE/Amex index was about 3.2 and its return volatility was about 20%.

Figure 5 plots the average realized returns (solid line) and expected returns (dashed line).[13] Ex post conditioning on the adoption of the new technology at time t^{**} generates

[13] All returns are annualized by multiplying each interval-dt return by $1/dt$.

stock returns that are first positive and then negative for both economies. This return pattern results from what we call a *hindsight bias*. Ex post, we know that a technological revolution took place at time t^{**}, but ex ante, we only have a probability assessment of this event. Before time t^{**}, stock prices are not expected to rise and fall; expected returns are given by the covariances with the stochastic discount factor (Corollary 2). However, conditioning on a technological revolution means that the adoption probability p_t must have been revised upward between times t^* and t^{**}, causing a bubble-like pattern in prices through the cash flow and discount rate effects discussed earlier. The bias of realized returns relative to expected returns is solely due to ex post conditioning on $p_{t^{**}} = 1$; when this conditioning is removed, the bias disappears. (Across all 50,000 simulations, average realized returns are equal to average expected returns.) The rise and fall in stock prices during technological revolutions are observable ex post but not predictable ex ante.

The unexpected arrival of the new technology causes the market value of the old economy to drop immediately, which is clear from the old economy's negative return at time t^* in Figure 5. This fall in market value is driven by two forces. First, the agent now anticipates that conversion costs might be paid at time t^{**}. Second, the possibility of eventual adoption increases systematic risk and so drives up the future discount rates.

4.1. Sensitivity Analysis

This section examines the sensitivity of the price dynamics to our parameter choices. Figure 6 is the counterpart of Panel A of Figure 4 (revolution), with various parameter changes.

In Panel A of Figure 6, risk aversion $\gamma = 4$, as opposed to $\gamma = 3$ in Figure 4. Higher risk aversion decreases M/B in both economies, as expected, but the pattern of M/B is otherwise the same as that in Figure 4. A hump-shaped pattern in M^N/B^N obtains for any $\gamma > 1$.

In Panel B of Figure 6, the cost of switching to the new technology is $\kappa = 0$, as opposed to $\kappa = 0.1$ in Figure 4. The only perceptible effect of the lower κ is to decrease M/B of the new economy. The reason is that the lower conversion cost makes it more likely that the new technology will be adopted, which increases discount rates and thus depresses prices. For the old economy, there is also a counterbalancing effect, as the lower conversion cost increases the old economy's post-conversion capital $B_{t^{**}_+} = B_{t^{**}_-}(1 - \kappa)$. The two effects approximately offset each other, so the old economy's M/B is almost unaffected by the change in κ. Most important, the price patterns look just like those in Figure 4.

In Panel C, prior uncertainty about ψ is $\widehat{\sigma}_{t^*} = 8\%$, compared to $\widehat{\sigma}_{t^*} = 4\%$ in Figure 4.

The higher uncertainty increases M^N/B^N, especially close to time t^* when p_t is small (equation (28)). However, as p_t increases during a revolution, uncertainty becomes increasingly systematic, pushing M^N/B^N down, and this discount rate effect is stronger when systematic uncertainty is higher. Therefore, in technological revolutions characterized by high uncertainty, the new economy firms tend to start out with high valuations that exhibit a large decline. High uncertainty amplifies the bubble-like pattern in stock prices.

In Panel D of Figure 6, the time until the adoption decision is shortened to $t^{**} - t^* = 4$, compared to $t^{**} - t^* = 8$ in Figure 4. Faster adoption increases M^N/B^N. To understand this effect, we note two facts. First, faster adoption implies higher uncertainty about ψ at time t^{**} because there is less time to learn (equation (10)). Second, faster adoption implies a higher adoption threshold $\underline{\psi}$ because t^{**} is lower and $\widehat{\sigma}_{t^{**}}$ is higher (equation (13)). Since $\widehat{\psi}_t$ has less time to reach a higher threshold, the adoption probability p_{t^*} is lower, which implies that systematic risk is initially lower and M^N/B^N starts higher than in Figure 4. M^N/B^N then rises higher and falls deeper than in Figure 4, conditional on $p_{t^{**}} = 1$, because both the cash flow effect and the discount rate effect are stronger when adoption is faster. The cash flow effect is stronger because in order for $\widehat{\psi}_t$ to reach a higher threshold in shorter time, the increase in $\widehat{\psi}_t$ must be sharper. The discount rate effect is stronger because uncertainty at time t^{**} is higher, and conditional on $p_{t^{**}} = 1$, this uncertainty is entirely systematic. Since both effects are stronger, the rise and fall in M^N/B^N are more striking than in Figure 4. Faster adoption of the new technology magnifies the bubble-like pattern in stock prices.

4.2. Optimal Adoption Time

In this section, we relax the assumption that t^{**} is exogenously given. Without this assumption, no closed-form solutions are available. We define the value function as

$$\mathcal{V}\left(B_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2, t; T\right) = E_t \left\{ \max_{t^{**}} E_{t^{**}} \left[\frac{W_T^{1-\gamma}}{1-\gamma} \right] \right\}, \tag{35}$$

where the maximization involves choosing the optimal time t^{**}, $t^* \leq t^{**} \leq T$ to adopt the new technology (no adoption, $t^{**} = T$, is a possibility). The agent has a real option to pay the conversion cost and adopt the new technology, and she solves for the best time to exercise this option. The value function in equation (35) satisfies a partial differential equation that we solve by using the finite difference method. The market prices and volatilities are also computed numerically. The details are in the Appendix.

Figure 7 plots the average paths of M/B and volatility when t^{**} is chosen optimally. Depending on the path of profitability, the adoption can occur anytime between t^* and T, but averaging across adoptions at very different t^{**}'s would not be very meaningful. For better comparison with Figure 4 in which t^{**} is fixed at 9 years, the left panels of Figure 7 report averages across those simulations in which the optimal t^{**} is between years 8 and 10. The right panels average across the simulations in which no revolution took place.

Figure 7 shows that our main results are unaffected by endogenizing t^{**}. The new economy's M/B is somewhat lower than in Figure 4, mostly because the optimal t^{**} exceeds 9 years, on average, and because slower adoption reduces M/B. More important, during revolutions, this M/B exhibits a rise-and-fall pattern similar to that in Figure 4. Realized returns in the new economy are positive at first and negative at last, due to the hindsight bias discussed earlier. The path of volatility in Panel C is also very similar to that in Figure 4. The main difference is that endogenous t^{**} produces a smoother path around time t^{**}: the volatility spike observed in Figure 4 disappears, as argued earlier.

5. American Railroads Before the Civil War

In this section we analyze the first major technological revolution that took place in the U.S. since the New York Stock Exchange was organized in 1792 – the introduction of steam-powered railroads (RRs). We argue that in the early days of the RR, there was substantial uncertainty about whether the RR technology would be ultimately adopted on a large scale. After examining the historical milestones of American RRs in Section 5.1., we argue that the probability of a large-scale adoption rose gradually, and that it approached one in the late 1850s after the RR expansion west of the Mississippi River. We then empirically examine the behavior of the RR stock prices in 1830–1861 in Section 5.2. In the context of our model, our evidence is consistent with large-scale adoption of the RR technology around year 1857.

5.1. Brief History

The steam engine, an 18th-century invention, was first used for rail-based transportation in the early 19th century in Britain. The United States followed shortly afterwards. The first RR act in the U.S. was passed in 1815 when the New Jersey legislature awarded a charter to Colonel John Stevens to build a RR between the Delaware and Raritan rivers.[14] In 1825, Stevens operated the first locomotive in America – his 16-foot "Steam Waggon" ran around a circular rail track in Hoboken at 12 miles per hour. The construction of the first RR, the Baltimore & Ohio, began in July 1828. The Baltimore & Ohio initially used

[14] The discussion in this section draws especially on Stover (1961), Fogel (1964), and Klein (1994).

horses to draw its cars, but it replaced them in 1830 by a steam locomotive, Peter Cooper's "Tom Thumb." In 1830, both passenger and freight service commenced on the Baltimore & Ohio. RRs spread quickly. On Christmas Day in 1830, the "Best Friend of Charleston," the first locomotive built for sale in the U.S., made the first scheduled steam-RR train run in America. Between 1830 and 1840, the RR mileage in the U.S. grew from 23 to 2,808 miles. In 1840, only four of the 26 states had not completed their first mile of track.

The new RR technology competed with the existing modes of transportation such as wagons, stagecoaches, steamboats, and canals. Those were not without problems – wagons were slow and expensive, stagecoaches were uncomfortable, steamboats were dangerous and limited in scope, and canals froze over in winter. However, it was far from obvious in the 1830s and 1840s that the RRs would later come to dominate the transportation industry. For example, waterways were much less expensive than RRs, and wagons were not restricted to rails. While the RR mileage caught up with the canal mileage in the early 1840s, waterways still carried the great bulk of the nation's freight in the late 1840s. Writes Fogel (1964): "Far from being viewed as essential to economic development, the first RRs were widely regarded as having only limited commercial application. Extreme skeptics argued that RRs were too crude to insure regular service, that the sparks thrown off by belching engines would set fire to buildings and fields, and that speeds of 20 to 30 miles per hour could be "fatal to wagons, road and loading, as well as to human life." More sober critics questioned the ability of RRs to provide low cost transportation, especially for heavy freight. [Some] placed "a RR between a good turnpike and a canal" in transportation efficiency."

Nearly all RRs organized as corporations funded by private investors. More than half of the more than $300 million invested in American RRs in 1850 was represented by capital stock, the remainder being in bonds. The freight business was economically more important than passenger traffic, which typically produced around 30% of the total revenue.

While most early RRs were built with local capital to provide local transportation, RR building became more ambitious in the 1850s. This decade "was one of the most dynamic periods in the history of American RRs" (Stover, 1961). RR mileage expanded from 9,021 in 1850 to 30,626 in 1860, and total investment in the industry increased from about $300m to about $1,150m over the same period. This growth was spurred by land grants to RRs by the federal government. The first land-granting act was passed by the Congress in 1850, aiding the Illinois Central and the Mobile & Ohio RRs. The RR growth in the 1850s was also stimulated by the discovery of gold in California and the lure of the trans-Pacific trade. In the 1850s, New York, Philadelphia, and Baltimore all achieved their rail connections with the west. In 1853, an all-rail route opened from the East to Chicago, and Chicago quickly became the rail capital of the nation. The RR technology also advanced in the 1850s – telegraph was first used to dispatch trains, T-rails became the general rule, and so did the standard track gauge, at least in the North.[15] "Instead of merely serving as connectors between navigable bodies of water as originally conceived, RRs were replacing them as the preferred way of transport" (Klein, 1994).

The dramatic RR growth in the 1850s is also evident in Figure 8, which plots the total rail consumption in the U.S., measured by the number of track-miles of rails laid each year (Fogel, 1964). Rail consumption grew fast in the 1830s, but especially fast during the decade leading up to 1856. After 1856, rail consumption slowed down and even declined in 1861 when the Civil War began, but it accelerated again after the war.

The diffusion of the RR technology made a leap in 1856 when two milestone RRs were completed: the Illinois Central, the longest RR in the world (705 miles), and the Sacramento Valley, the first RR in California. Also in 1856, the first RR bridge across the Mississippi was built near Davenport, Iowa, heralding future westward expansion into the region then known as the "Great American Desert." This westward expansion was the defining feature of the RR growth in the decades to come. The RRs shaped the economy of the West, creating new national markets and fostering unprecedented economic specialization across the nation.

By the late 1850s, it seemed clear that the RR had become a dominant form of transportation. According to Stover (1961), "By 1860 the canal packets and river steamers had lost much of their passenger traffic" to the RR. In 1860, every state save Minnesota and Oregon had RR mileage, and 29 of the 33 states had more than 100 miles of line. Klein (1994) argues that "By 1860... [the RR] had emerged not only as the preferred form of transportation but also as the chief weapon of commercial rivalry." This evidence suggests that a large-scale adoption of the RR technology took place by the end of the 1850s.

5.2. Railroad Stock Prices

To examine the behavior of RR stock prices in the early days of the RR (1830–1861), we use the data compiled by Goetzmann, Ibbotson, and Peng (2001). These data contain monthly individual stock prices for NYSE stocks from 1815 to 1925, as well as annual dividends for

[15]The Northern RRs were using 11 different track gauges in the 1850s, but the standard gauge, 4'8.5", became by far the most common by 1860, according to Stover (1961). The South was still mostly on the 5' gauge. Benmelech (2005) exploits the diversity of track gauges in 19th century American railroads to examine the effect of asset liquidation value on capital structure.

a subset of stocks from 1825 to 1870. The data are provided by the International Center of Finance at Yale University at http://icf.som.yale.edu/nyse/ (as of January 7, 2005).

To focus on common stocks, we exclude stocks classified as "preferred" or "scrip" in the database. (Scrips are certificates convertible into shares when fully paid-in.) If such classification is not provided, we examine the stock name and exclude stocks whose name contains an indication of non-common status such as "pref," "pr.," "pf," or "scrip." Among the 671 stocks in the database, we identify and exclude 85 preferred stocks and 29 scrips.

We identify RR stocks by examining the stock names. As an initial benchmark, we use the list of RRs from the DeGolyer Library at Southern Methodist University, provided at http://www.smu.edu/cul/degolyer/rr%20names%20for%20web.htm. This list is comprehensive but incomplete. When in doubt, we search the internet for more information. Overall, we identify 284 RR stocks (42.32% of the whole sample). The first RRs that appear in our price index (discussed below) in 1831 are Camden & Amboy, Canajoharie & Catskill, Harlem, and Ithaca & Oswego. All RRs that have at least one valid monthly common stock return between 1830 and 1861 are listed in Table 2.

We clean the monthly price file to remove apparent data errors. To proceed in a systematic fashion, we exclude all prices that imply implausibly large return reversals. Specifically, we exclude prices that more than tripled compared to the most recent available price and then fell to less than a third at the nearest future observation, as well as prices that experienced the same reversals in reverse order (first down, then up). We eliminate 34 such prices in our 1830–1861 sample. We also examine all price sequences in which the price increased or decreased at least tenfold without reversal, and eliminate six suspicious price entries between 1830 and 1861. We retain the price entries that imply returns below -90% at the very end of a stock's price series because these could be stocks heading for bankruptcy. Altogether, we delete 40 of the 15,276 price entries between 1830 and 1861, or 0.26% of the sample.

Before the price coverage in the database improves in 1848, uninterrupted price sequences for RR stocks are rare. In no month before 1848 are there more than five RR stocks with valid monthly returns, and there are months with zero RR returns. An important part of the problem are gaps in the price series, in which one or several missing values are sandwiched between two valid prices for a given stock. To alleviate the data shortage, we fill in such gaps by linear interpolation, but only for gaps that are no more than three months long. This procedure substantially increases the price coverage early in the sample. For example, without interpolating, the RR year-end price-dividend ratio discussed below would have only three valid observations prior to 1847; with interpolation, the number of valid observations increases to eight. Without interpolating, our results would be noisier, with more missing values, but they would lead to the same basic conclusions.

Panel A of Figure 9 plots the aggregate price-to-dividend ratio (P/D) for the RR and non-RR industries. Each year, we compute P/D as the sum of year-end prices divided by the sum of dividends paid in that year, summing across all RR (or non-RR) stocks with valid price and dividend data. Note three main results. First, the P/D of RRs almost invariably exceeds the P/D of non-RRs before the mid-1850s. Second, the RR P/D falls from 24.9 in 1846 to 15.8 in 1852, to 6.5 in 1857. Third, the non-RR P/D falls as well, but less dramatically: from 14.0 to 12.8 to 9.1 over the same period. While interpreting the noisy data requires caution, all three results are broadly consistent with the idea that the new RR technology was widely adopted around 1857.

Panel B of Figure 9 plots the stock price indexes for the RR and non-RR industries, obtained by cumulating monthly returns in each industry. Industry returns are computed as price-weighted averages of monthly capital gains across all stocks in the industry.[16] We use capital gains rather than total returns because the dividend data available to us are annual, not monthly, and because these data are spotty, especially early in the sample (Goetzmann et al. (2001) suggest that their dividend sample is incomplete). The general downward trend in the price indexes is partly due to the absence of dividends and partly due to the absence of inflation in the economy. The biggest price declines occur in the mid-1850s. For example, between June 1853 and October 1857, the RR price index falls by 58.3%, whereas the non-RR index falls by 33.9%. Both the sharp price decline for RRs and the milder decline for non-RRs are consistent with the RR technology being adopted on a large scale around 1857. Recall that our model predicts that the new economy (RR) stock prices fall by more than the old economy (non-RR) stock prices before the adoption of the new technology.

Various events played a role in the stock price decline in 1857. Investor confidence was shaken by embezzlement at the Ohio Life Insurance and Trust Company in August, as well as by the government's loss of a large amount of gold at sea in September. Other commonly cited negative influences include falling grain prices, British withdrawals of capital from U.S. banks, and manufacturing surpluses. The stock market bottomed in October 1857 amidst a number of bank failures. However, the stock price decline cannot be fully attributed to the banking panic. According to Mishkin (1991), "Rather than starting with the banking panic in October 1857, the disturbance to the financial markets seems to arise several months

[16]Goetzmann et al. (2001) argue that price-weighting best approximates the return on a buy-and-hold portfolio, given the absence of information about market capitalization and book value in their database.

earlier with the rise in interest rates, the stock market decline... and the widening of the interest rate spread." Mishkins last observation is particularly interesting. He shows that the spread between the yields of low- and high-quality corporate bonds was unusually high in 1857–1859, higher than at any future time before the 1930s. These high yield spreads indicate that the risk premia in the late 1850s were high, consistent with our story. Mishkin also opines that the decline in stock prices in the late 1850s "might be linked to the general rise in interest rates which lowers the present discounted value of future income streams." This is precisely our story - stock prices fall shortly before the adoption of the new technology because discount rates increase due to an increase in systematic risk.

Panel C plots the volatility of returns in the RR and non-RR industries, computed annually as the standard deviation of monthly industry returns within the year. Two facts seem noteworthy. First, the RR volatility exceeds the non-RR volatility in every year except 1841, consistent with the presence of uncertainty about the RR technology. This fact is also consistent with the explanation that the RR portfolio is less diversified than the non-RR portfolio, which is apparent from Panel D. While we cannot dismiss this alternative explanation, we note that the volatility difference persists also after the number of RRs with valid monthly stock returns increases sharply (from 6 in December 1847 to 15 in January 1848, to 25 in July 1850). The second interesting fact in Panel C is that return volatility increases sharply in 1857, to 33.5% per year for RRs and to 23.1% for non-RRs. Again, this fact is consistent with a large-scale adoption of the RR technology around 1857.

6. Conclusions

We provide a rational explanation for the bubble-like patterns in stock prices observed during technological revolutions. Stock prices of innovative firms initially rise due to good news about the productivity of the new technology, but they ultimately fall as the risk of the technology changes from idiosyncratic to systematic. The rise and fall in prices are observable only in hindsight. This price pattern is unexpected while investors are uncertain whether the new technology would be widely adopted, but we observe it ex post because we focus on technologies that eventually led to technological revolutions. To formalize this intuition, we develop and calibrate a general equilibrium model that features a real option decision and Bayesian learning about the average productivity of the new technology.

According to the model, the bubble-like price pattern should be most pronounced for technologies characterized by high uncertainty and fast adoption. These characteristics seem to fit the Internet technology, which spread quickly over the past decade amid high uncertainty about the future growth of the Web. The recent "bubble" in the Internet stock prices might be to some extent due to the mechanism discussed here.[17]

Although we focus on stock prices, our model also has some implications for productivity. First, the new technology does not bring productivity gains immediately upon arrival because the agent finds it optimal to learn about a new technology before adopting it. Since the agent chooses the adoption time optimally depending on what she learns, the time it takes for the productivity gains to begin emerging is endogenous in the model. The implication that productivity gains arrive with a lag seems reasonable; for example, although electric power first appeared around 1880, it was not until the 1920s that the productivity of the U.S. economy increased as a result of a large-scale adoption of electricity (David, 1991). Another realistic implication is that productivity increases gradually after the adoption of the new technology. Although the adoption occurs suddenly, productivity rises slowly instead of jumping because the adoption shifts the long-run mean of productivity.

Our model has no direct implications for investment. The agent invests only a negligible amount in the new technology for learning purposes. Investing more would not allow the agent to learn faster because there is only one stream of signals about the productivity gain (the new economy's realized productivity) and any investment in the new economy allows the agent to observe this signal. In an extension that would allow multiple or costly signals, the amount invested could affect the speed of learning.[18] Such an extension might have novel implications for investment while preserving the pricing implications of our model.

Economists typically date technological revolutions by analyzing the underlying inventions, productivity growth, or product prices. We offer a complementary indirect approach based on the behavior of stock prices. To give an example, we examine stock prices during the 19th century railroad revolution in the United States, and find evidence consistent with a large-scale adoption of the railroad technology soon after railroads began expanding west of the Mississippi. A systematic empirical study of stock prices during technological revolutions is beyond the scope of this paper, but it is a promising avenue for future research.

[17] "The growth of the Internet has paralleled that of most industries based on revolutionary technology. Canals, railroads, telegraphs, telephones, cars, radios, personal computers – all progressed (or are progressing) through four phases of development: boom, bust, mature growth and decay... the repetition of the pattern... suggests that the boom-and-bust phases should be viewed as far more than repeated examples of human folly." (The New York Times, "Irreplaceable Exuberance," August 30, 2005.)

[18] A similar mechanism is at work in the model of Johnson (2005) who argues that learning about the curvature of the production function of a new technology can generate overinvestment in this technology.

Table 1

Parameters used in Simulations.

	$\overline{\rho}$	$\widehat{\psi}_{t^*}$	$\widehat{\sigma}_{t^*}$
	0.1217	0	0.04
ϕ	σ_0	$\sigma_{N,0}$	$\sigma_{N,1}$
0.3551	0.07	0.07	0.07
κ	$t^{**} - t^*$	T	γ
0.1	8	30	3

Table 2
Railroads Appearing in our Price Index.

This table lists all railroads in our sample that have at least one valid monthly common stock return between 1830 and 1861. The railroads are sorted by the year of appearance of their first valid monthly return.

Year	Railroad
1831	Camden & Amboy; Canajoharie & Catskill; Harlem; Ithaca & Oswego
1832	Boston & Providence
1833	Boston & Worcester; Brooklyn & Jamaica
1835	Hudson & Berkshire; Long Island
1839	Auburn & Syracuse
1841	Auburn & Rochester
1844	Housatonic
1847	Hudson River; Macon & West
1848	Hartford & New Haven; New York & Erie
1849	Erie
1850	Albany & Schenectady; Baltimore & Ohio; Michigan Central; New York & Harlem
1851	Chemung
1852	Michigan & Southern
1853	Cincinnati, Hamilton & Dayton; Cleveland, Columbus & Cincinnati; Cleveland & Pittsburg; Cleveland & Toledo; Galena & Chicago; Illinois Central; Little Miami
1854	Chicago & Rock Island
1855	Michigan Southern & Northern Indiana
1856	Eighth Avenue; Lacrosse & Milwaukee; Macon & Western
1857	Chicago, Burlington & Quincy; Delaware, Lackawanna & Western; Indianapolis & Cincinnati
1858	Brooklyn City; Buffalo & State Line; Cleveland, Painesville & Ashtabula

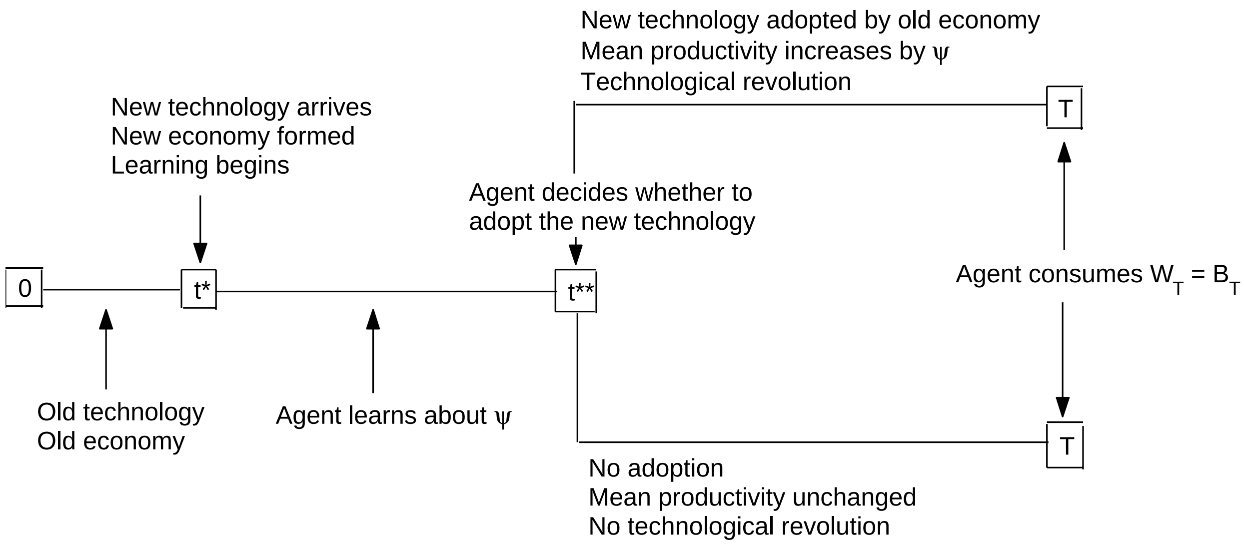

Figure 1. The Sequence of Events. In this chart, t^{**}, the time when the agent decides whether to adopt the new technology, is taken as given. We initially take t^{**} as given for the purpose of obtaining closed-form solutions for prices, but later we solve for the optimal time t^{**} to adopt the new technology.

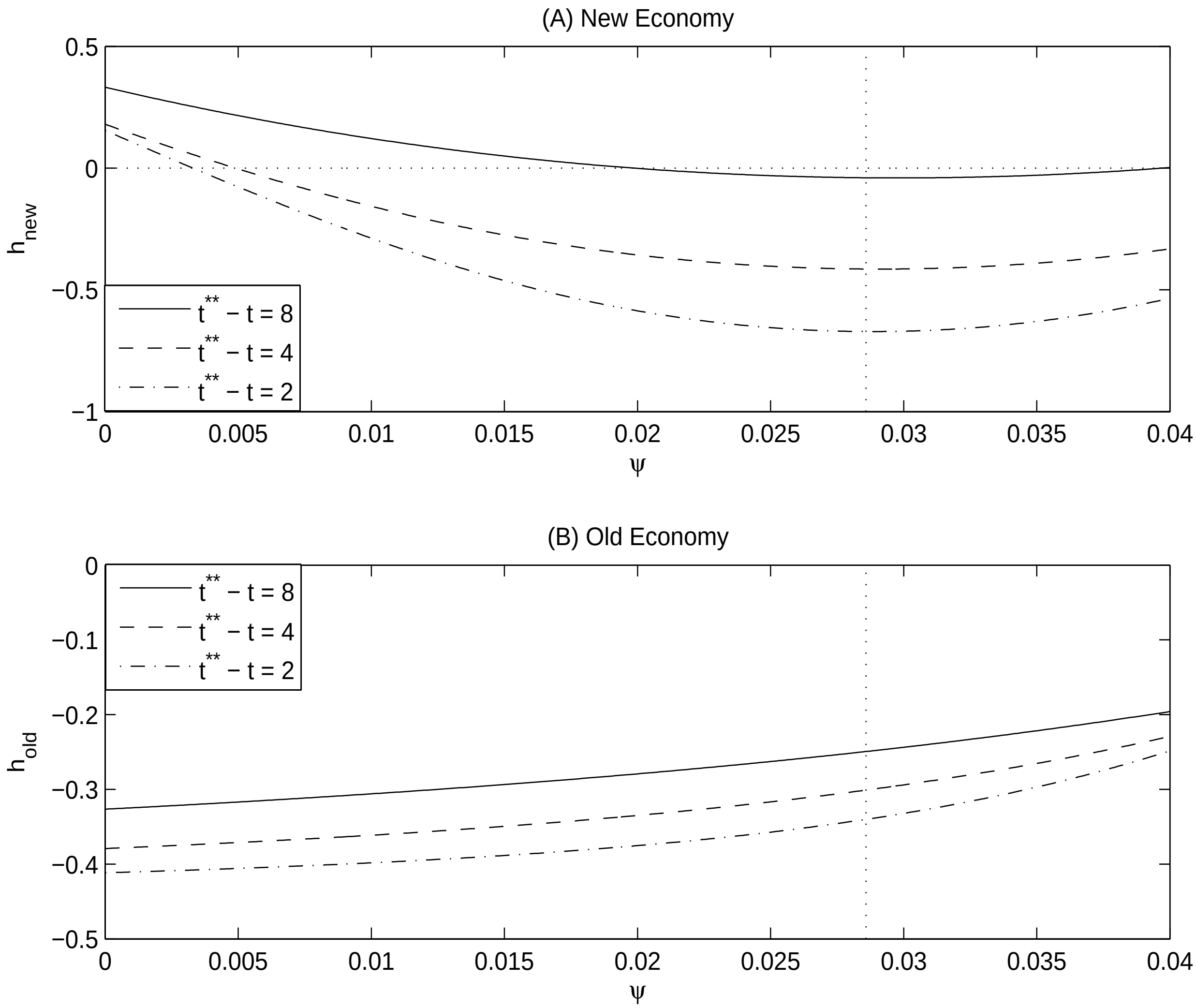

Figure 2. Functions h_{new} and h_{old} from Proposition 6. This figure plots h_{new} (Panel A) and h_{old} (Panel B) as functions of $\widehat{\psi}_t$ and $t^{**} - t$, for parameter values in Table 1. The dotted vertical lines mark $\underline{\psi}$.

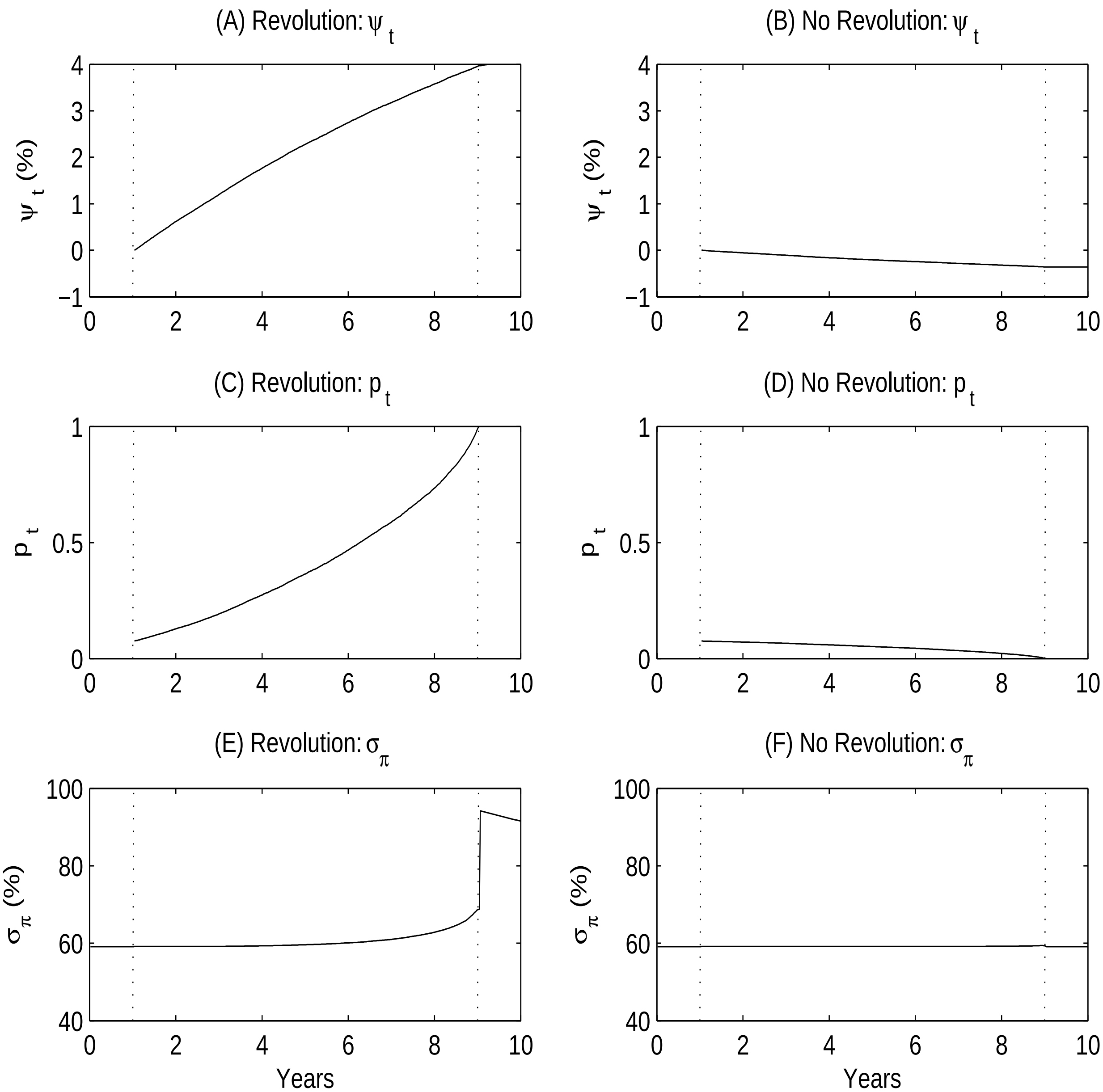

Figure 3. Average $\widehat{\psi}_t$, p_t, $\sigma_{\pi,t}$ in Simulations. The left panels plot the perceived productivity gain $\widehat{\psi}_t$ (Panel A), the adoption probability p_t (Panel C), and the volatility of the stochastic discount factor $\sigma_{\pi,t}$ (Panel E), averaged across all simulations in which the new technology was adopted at time t^{**} ($p_{t^{**}} = 1$). The right panels (B, D, and E) plot the same quantities but the average is taken across all simulations in which the new technology was not adopted at time t^{**} ($p_{t^{**}} = 0$). In each panel, the first vertical line denotes $t^* = 1$, the time when the new technology becomes available, and the second vertical line denotes $t^{**} = 9$, the time at which the agent decides whether to adopt the technology on a large scale. All parameters are given in Table 1.

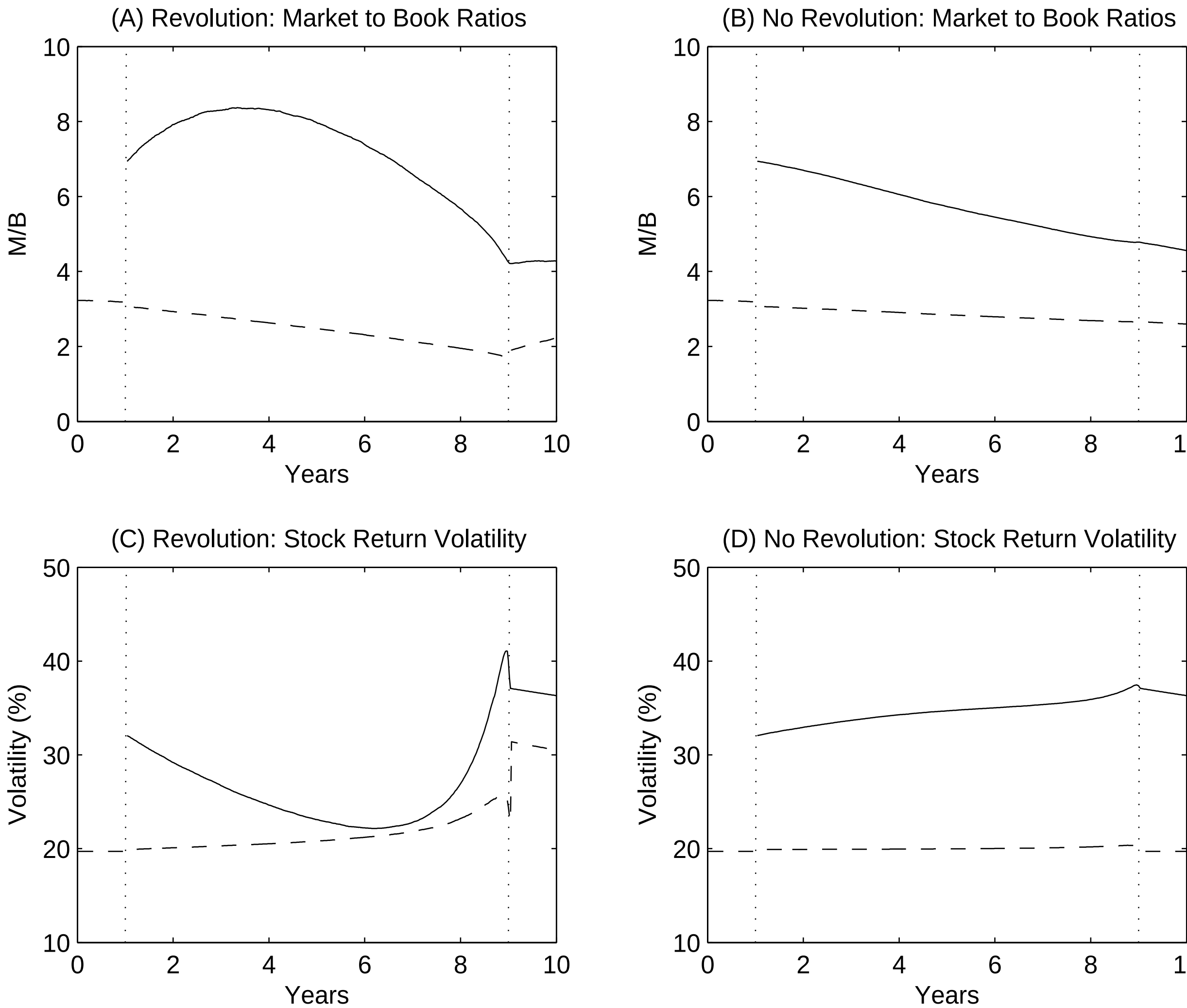

Figure 4. Average M/B and Volatility in Simulations. Panel A plots the path of the market-to-book ratio of the new economy (solid line) and old economy (dashed line) averaged across all simulations in which the new technology was adopted at time t^{**} ($p_{t^{**}} = 1$). Panel B is an equivalent of Panel A, except that the averages are computed across all simulations in which the new technology was not adopted at time t^{**} ($p_{t^{**}} = 0$). Panels C and D are equivalents of Panels A and B, respectively, with M/B replaced by the volatility of stock returns. In each panel, the first vertical line denotes $t^* = 1$, the time when the new technology becomes available, and the second vertical line denotes $t^{**} = 9$, the time at which the agent decides whether to adopt the technology on a large scale. All parameters are given in Table 1.

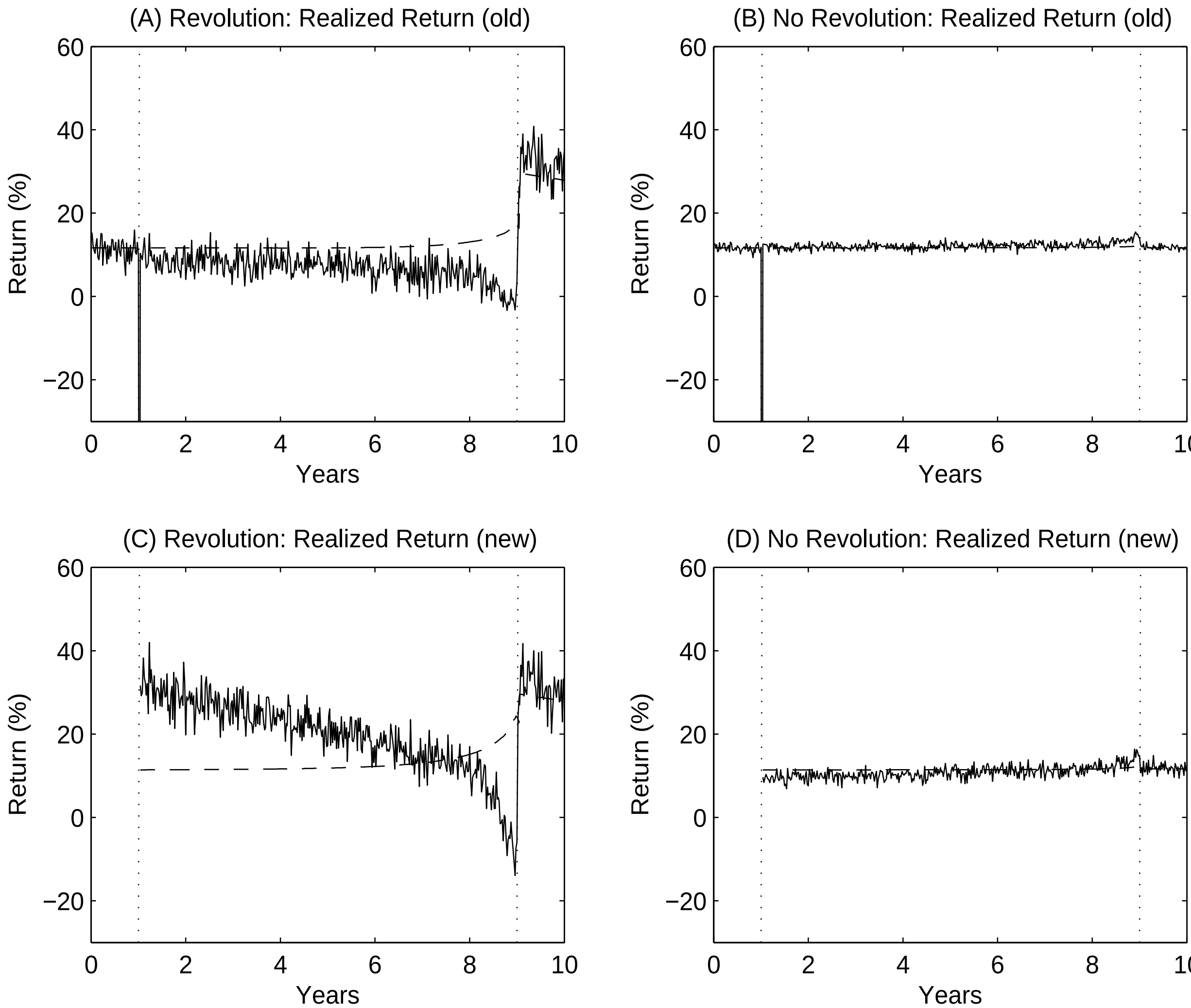

Figure 5. Average Realized and Expected Stock Return in Simulations. The left panels plot the realized return (solid line) and expected return (dashed line) for the old economy (Panel A) and the new economy (Panel C), averaged across all simulations in which the new technology was adopted at time t^{**} ($p_{t^{**}} = 1$). The right panels (B and D) plot the same quantities but the average is taken across all simulations in which the new technology was not adopted at time t^{**} ($p_{t^{**}} = 0$). In each panel, the first vertical line denotes $t^* = 1$, the time when the new technology becomes available, and the second vertical line denotes $t^{**} = 9$, the time at which the agent decides whether to adopt the technology on a large scale. All parameters are given in Table 1.

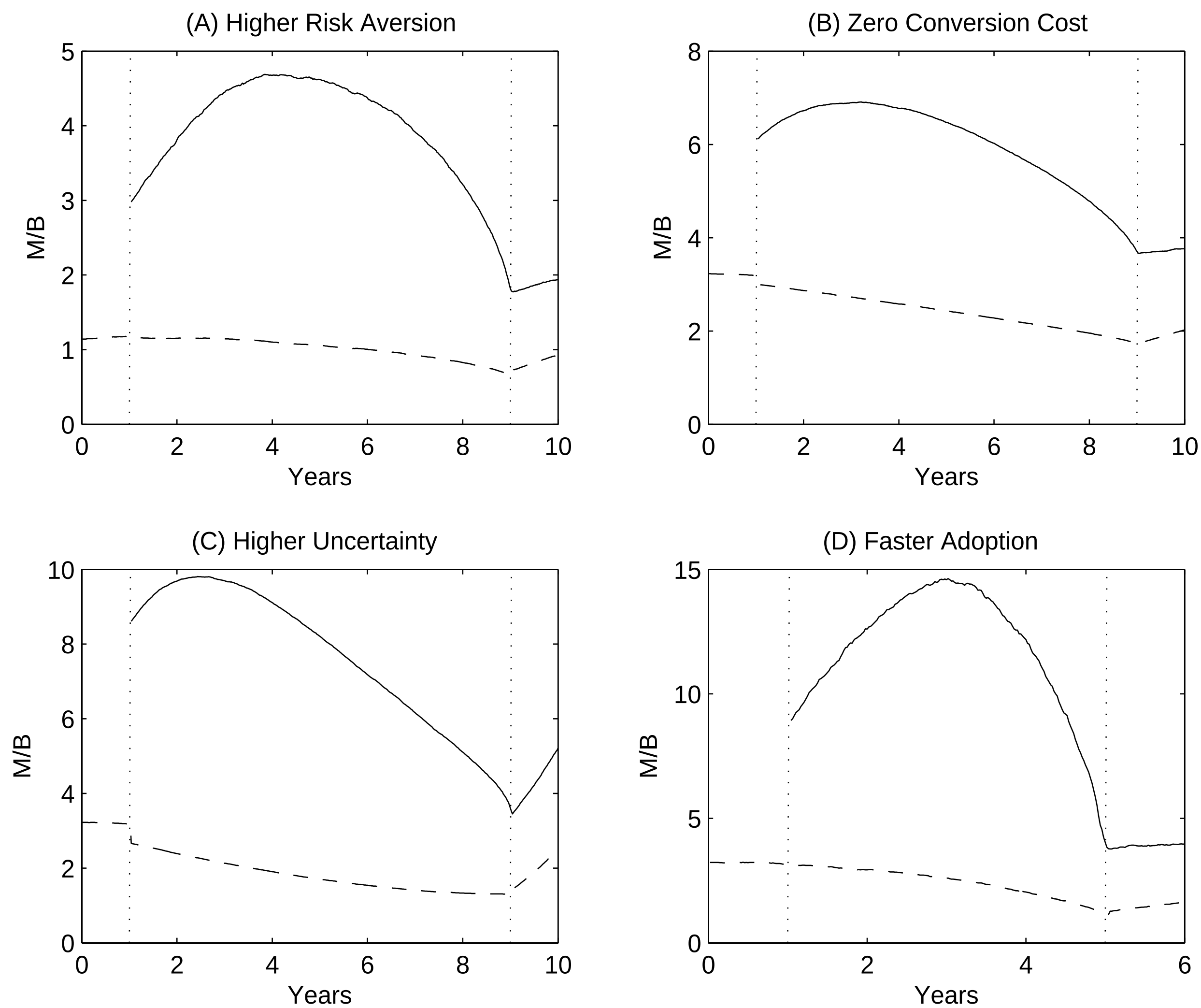

Figure 6. Average M/B in Simulated Revolutions: Sensitivity Analysis. All four panels plot the paths of the market-to-book ratio of the new economy (solid line) and old economy (dashed line) averaged across all simulations in which the new technology was adopted at time t^{**} ($p_{t^{**}} = 1$). All parameters are given in Table 1, except for one change that varies across the panels. In Panel A, the risk aversion $\gamma = 4$ instead of the benchmark case $\gamma = 3$. In Panel B, the conversion cost $\kappa = 0$ instead of the benchmark case $\kappa = 0.1$. In Panel C, the uncertainty $\sigma_{t^*} = 0.08$ instead of the benchmark case $\sigma_{t^*} = 0.04$. In Panel D, the time until the adoption $t^{**} - t^* = 4$ instead of the benchmark case $t^{**} - t^* = 8$ years. In each panel, the first vertical line denotes t^*, the time when the new technology becomes available, and the second vertical line denotes t^{**}, the time at which the agent decides whether to adopt the technology on a large scale.

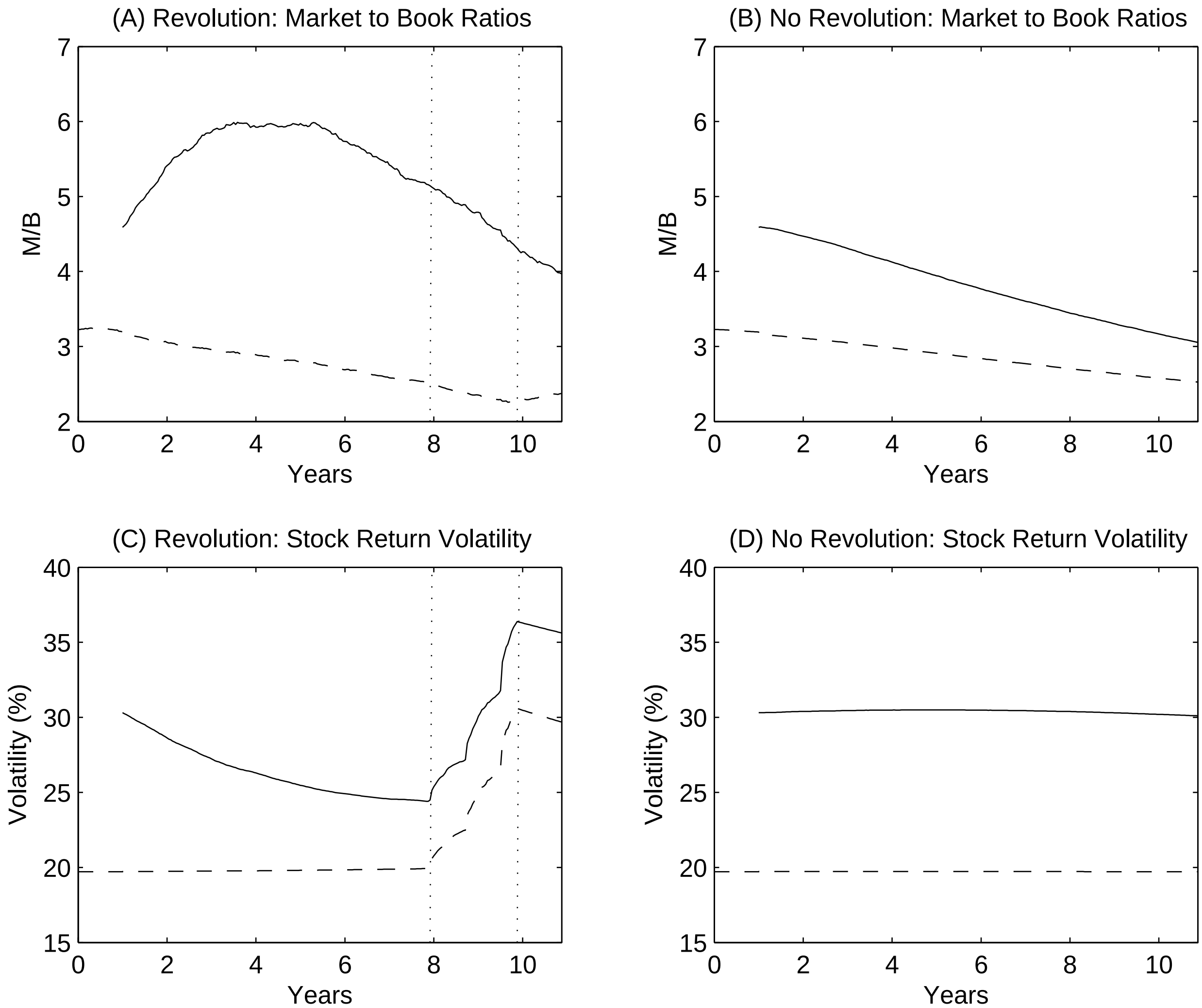

Figure 7. Average M/B and Volatility in Simulations with Optimal Adoption Time. Panel A plots the path of the market-to-book ratio of the new economy (solid line) and old economy (dashed line) averaged across all simulations in which the new technology was adopted at an optimally chosen time t^{**} between years 8 and 10. Panel B is an equivalent of Panel A, except that the averages are computed across all simulations in which the new technology was never adopted. Panels C and D are equivalents of Panels A and B, respectively, with M/B replaced by the volatility of stock returns. All parameters are in Table 1.

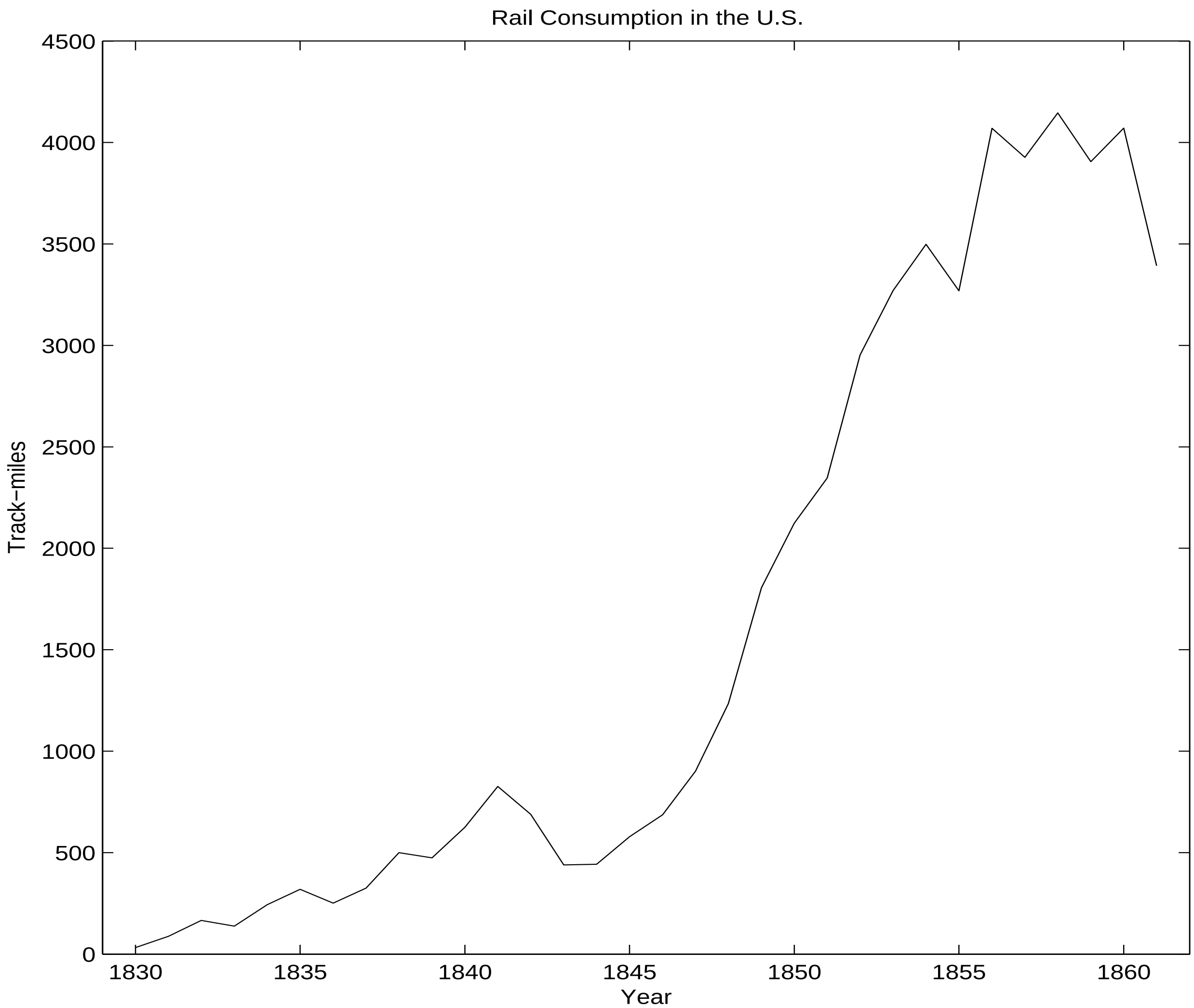

Figure 8. Total Rail Consumption in the United States. The figure plots the number of track-miles of rails laid each year in the U.S., as estimated by Fogel (1964, p.174). A track-mile of rails is defined as one half of the length of the rails in a mile of single track. The total includes rails used in the construction of new track as well as in the replacement of worn-out rails.

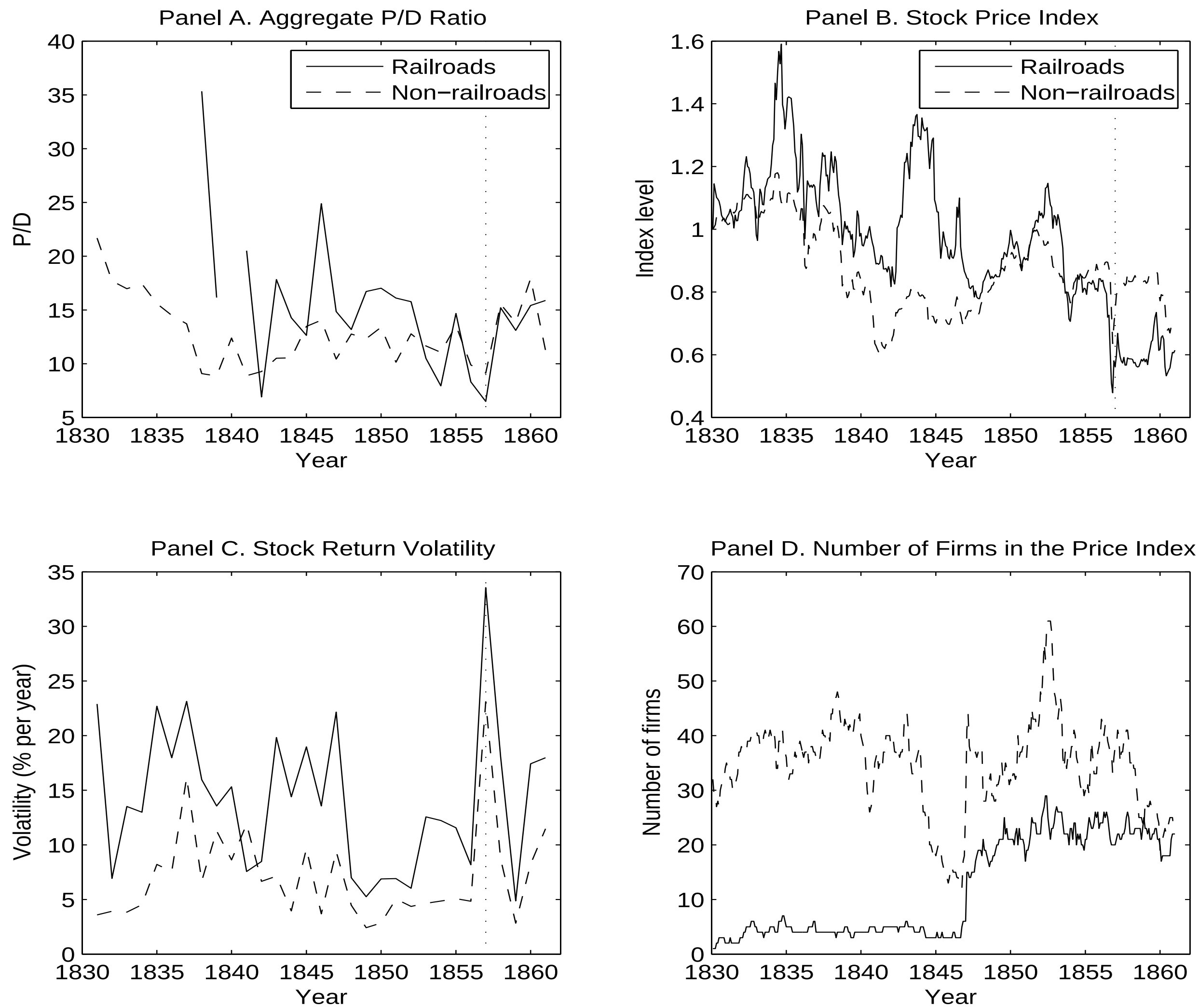

Figure 9. Railroad and Non-Railroad Common Stock Prices. Panel A plots the aggregate price-to-dividend ratio for the railroad (solid line) and non-railroad (dashed line) industries. Each year, this ratio is computed as the sum of year-end prices divided by the sum of dividends paid in that year, summing across all railroad (non-railroad) stocks with valid price and dividend data. Panel B plots the stock price index, obtained by cumulating monthly capital gains in the respective industry. Each month, the capital gain for railroads (non-railroads) is computed as the price-weighted average of monthly capital gains across all railroad (non-railroad) stocks. Panel C plots the standard deviation of returns in the railroad and non-railroad industries. Each year, this standard deviation is computed across all monthly price-weighted average industry returns in the given year. Panel D plots the number of firms with valid monthly stock returns in each month.

Appendix.

The Appendix contains the sketches of all proofs. The formal proofs are available in the companion Technical Appendix, which is downloadable from the authors' websites.

Lemma A1: Let $\tau = T - t$. The expectation in equation (6) is given by

$$V\left(B_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2, t; T\right) = E_t\left[\frac{B_T^{1-\gamma}}{1-\gamma}\right] = \frac{B_t^{1-\gamma}}{1-\gamma} e^{A_0(\tau)+(1-\gamma)A_1(\tau)\rho_t+(1-\gamma)A_2(\tau)\widehat{\psi}_t+\frac{1}{2}(1-\gamma)^2 A_2(\tau)^2\widehat{\sigma}_t^2}, \tag{36}$$

where $A_1(\tau)$ and $A_2(\tau)$ are given in Propositions 4 and 2, respectively, and

$$A_0(\tau) = (1-\gamma)\overline{\rho}(\tau - A_1(\tau)) + \frac{\sigma^2}{2}\frac{(1-\gamma)^2}{\phi^2}\left\{\tau + \frac{1-e^{-2\phi\tau}}{2\phi} - 2\frac{1-e^{-\phi\tau}}{\phi}\right\}.$$

Proof: Let $\mathbf{x}_t = \left(b_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2\right)$. From the Feynman-Kac theorem, V satisfies

$$0 = \frac{\partial V}{\partial t} + \sum_i \frac{\partial V}{\partial x_i} E_t[dx_i] + \frac{1}{2}\sum_i\sum_j \frac{\partial^2 V}{\partial x_i \partial x_j} E_t[dx_i dx_j],$$

with the boundary condition $V(\mathbf{x}_T) = (1-\gamma)^{-1} e^{(1-\gamma)x_{1,T}}$. This PDE is satisfied by (36).

Proof of Proposition 1. Since $\gamma > 1$, V in equation (36) is negative, decreasing in $\widehat{\sigma}_t^2$, and increasing in B_t. As a result, $V\left(B_{t^*}(1-\kappa), \rho_{t^*}, 0, \widehat{\sigma}_{t^*}^2, t^*; T\right) < V\left(B_{t^*}, \rho_{t^*}, 0, 0, t^*; T\right).$

Proof of Lemma 1. Given the observation equations (3) and (8), the result follows from Theorem 10.3 in Liptser and Shiryayev (1977). The explicit formulas for $(\widetilde{Z}_{0,t}, \widetilde{Z}_{1,t})$, which capture perceived innovations in ρ_t^N and ρ_t, are given in the Technical Appendix.

Proof of Proposition 2. Using Lemma A1, it is easy to verify that (13) follows from (14).

Lemma A2: The distribution of $\widehat{\psi}_{t^{**}}$ conditional on $\widehat{\psi}_t$ is normal:

$$\widehat{\psi}_{t^{**}}|_{\widehat{\psi}_t} \sim N\left(\widehat{\psi}_t, \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right).$$

In addition, $p_t \equiv \text{Prob}\left(\widehat{\psi}_{t^{**}} > \underline{\psi}|\widehat{\psi}_t\right) = 1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right).$

Proof: The process for $\widehat{\psi}_t$ is linear with deterministic volatility. The result then follows.

Lemma A3: For $t^* \leq t < t^{**}$, the value function in equation (15) is given by

$$\mathcal{V}\left(B_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2, t; T\right) = \frac{B_t^{1-\gamma}}{1-\gamma}\left\{(1-p_t)G_t^{no} + p_t G_t^{yes}\right\}, \tag{37}$$

where

$$G_t^{no} = e^{A_0(\tau)+(1-\gamma)A_1(\tau)\rho_t} \tag{38}$$

$$G_t^{yes} = G_t^{no}(1-\kappa)^{1-\gamma} R_t e^{(1-\gamma)A_2(\tau^{**})\widehat{\psi}_t+\frac{1}{2}(1-\gamma)^2 A_2(\tau^{**})^2\widehat{\sigma}_t^2} \tag{39}$$

and

$$R_t = \frac{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + (1-\gamma) A_2(\tau^{**}) \left(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right), \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)}{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)} < 1. \tag{40}$$

Proof: From the definition of the value function and $W_T = B_T$, we have

$$\mathcal{V}\left(B_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2, t; T\right) = (1 - p_t) E_t \left[\frac{B_T^{1-\gamma}}{1-\gamma} | \widehat{\psi}_{t^{**}} < \underline{\psi}\right] + p_t E_t \left[\frac{B_T^{1-\gamma}}{1-\gamma} | \widehat{\psi}_{t^{**}} \geq \underline{\psi}\right],$$

as the adoption occurs at t^{**} if and only if $\widehat{\psi}_{t^{**}} \geq \underline{\psi}$. Explicit computations show that

$$E_t \left[\frac{B_T^{1-\gamma}}{1-\gamma} | \widehat{\psi}_{t^{**}} < \underline{\psi}\right] = \frac{B_t^{1-\gamma}}{1-\gamma} G_t^{no} \quad \text{and} \quad E_t \left[\frac{B_T^{1-\gamma}}{1-\gamma} | \widehat{\psi}_{t^{**}} \geq \underline{\psi}\right] = \frac{B_t^{1-\gamma}}{1-\gamma} G_t^{yes}.$$

Proof of Proposition 3. From Lemma A1, $V\left(B_{t^*}, \rho_{t^*}, 0, 0, t; T^*\right) = B_{t^*}^{1-\gamma} / (1-\gamma)\, G_{t^*}^{no}$. Comparing this formula with $\mathcal{V}\left(B_{t^*}, \rho_{t^*}, 0, \widehat{\sigma}_{t^*}^2, t^*; T\right)$ and recalling that $\gamma > 1$, claim (16) follows if $G_t^{yes} < G_t^{no}$. The fraction G_t^{yes}/G_t^{no} can be shown to equal J_t, which is given by

$$J_t = E_t \left[e^{(1-\gamma)\log(1-\kappa) + (1-\gamma) A_2(t^{**};T) \widehat{\psi}_{t^{**}} + \frac{1}{2}(1-\gamma)^2 A_2(t^{**};T)^2 \widehat{\sigma}_{t^{**}}^2} | \widehat{\psi}_{t^{**}} > \underline{\psi}\right].$$

Using the definition of $\underline{\psi}$ in equation (13), J_t can be rewritten as

$$J_t = E_t \left[e^{(1-\gamma) A_2(\tau^{**})\left[\widehat{\psi}_{t^{**}} - \underline{\psi}\right]} | \widehat{\psi}_{t^{**}} > \underline{\psi}\right].$$

Since J_t is an expectation of a random variable that is always less than 1, we have $J_t < 1$.

Proof of Proposition 4. The proof is analogous to that of Lemma A3, except that "$(1-\gamma)$" is substituted with "$-\gamma$". Explicit calculations show that

$$\widetilde{G}_t^{yes} \equiv E\left[\left(\frac{B_T}{B_t}\right)^{-\gamma} | \widehat{\psi}_{t^{**}} \geq \underline{\psi}\right] = \widetilde{G}_t^{no} (1-\kappa)^{-\gamma}\, \widetilde{R}_t e^{-\gamma A_2(\tau^{**}) \widehat{\psi}_t + \frac{1}{2}\gamma^2 A_2(\tau^{**})^2 \widehat{\sigma}_t^2}, \tag{41}$$

where $\widetilde{G}_t^{no}$ is given in equation (21) and

$$\widetilde{R}_t = \frac{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2(\tau^{**}) \left(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right), \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)}{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)} < 1 \tag{42}$$

$$\overline{A}_0(\tau) = -\gamma \overline{\rho}\left(\tau - A_1(\tau)\right) + \frac{\sigma^2}{2}\frac{\gamma^2}{\phi^2}\left\{\tau + \frac{1 - e^{-2\phi\tau}}{2\phi} - 2\frac{1 - e^{-\phi\tau}}{\phi}\right\}. \tag{43}$$

Proof of Corollary 1. Let $\widetilde{p}_t = 1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2(\tau^{**}) \left(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right), \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)$. The claim follows from an application of Ito's Lemma, where

$$S_{\pi,t} = \frac{\left(\gamma A_2(\tau^{**}) - \frac{1}{\widetilde{p}_t}\frac{\partial \widetilde{p}_t}{\partial \widehat{\psi}_t}\right) \widetilde{G}_t^{yes} + \frac{\partial p_t}{\partial \widehat{\psi}_t} \widetilde{G}_t^{no}}{(1-p_t)\, \widetilde{G}_t^{no} + p_t \widetilde{G}_t^{yes}}. \tag{44}$$

Proof of Proposition 5. The old economy result follows from $M_t = E_t[\pi_T B_T]/\pi_t = E_t\left[B_T^{1-\gamma}\right]/E_t\left[B_T^{-\gamma}\right]$, as well as from Lemma A3 and Proposition 4. For the new economy, explicit computations of the conditional expectations show that

$$\begin{aligned} K_t^{no} &\equiv E_t\left[\left(\frac{B_T}{B_t}\right)^{-\gamma}\frac{B_T^N}{B_t^N}|\widehat{\psi}_{t^{**}}<\underline{\psi}\right] = K_t R_{L,t}^N \\ K_t^{yes} &\equiv E_t\left[\left(\frac{B_T}{B_t}\right)^{-\gamma}\frac{B_T^N}{B_t^N}|\widehat{\psi}_{t^{**}}\geq\underline{\psi}\right] = (1-\kappa)^{-\gamma} K_t^N R_{H,t}^N, \end{aligned}$$

where

$$\begin{aligned} K_t &= e^{C_0(\tau)-\gamma A_1(\tau)\rho_t + A_1(\tau)\rho_t^N + A_2(\tau)\widehat{\psi}_t + \frac{1}{2}A_2^2(\tau)\widehat{\sigma}_t^2} \\ K_t^N &= K_t e^{-\gamma A_2(\tau^{**})\widehat{\psi}_t + \frac{1}{2}\gamma A_2(\tau^{**})(\gamma A_2(\tau^{**})-2A_2(\tau))\widehat{\sigma}_t^2} \end{aligned}$$

and

$$\begin{aligned} R_{L,t}^N &= \frac{\mathcal{N}\left(\underline{\psi};\widehat{\psi}_t+\sigma_{y\widehat{\psi}}^L,\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)}{\mathcal{N}\left(\underline{\psi};\widehat{\psi}_t,\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)} \text{ with } \sigma_{y\widehat{\psi}}^L = A_2(\tau)\,\widehat{\sigma}_t^2 - A_2(\tau^{**})\,\widehat{\sigma}_{t^{**}}^2 \\ R_{H,t}^N &= \frac{1-\mathcal{N}\left(\underline{\psi};\widehat{\psi}_t+\sigma_{y\widehat{\psi}}^H,\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)}{1-\mathcal{N}\left(\underline{\psi};\widehat{\psi}_t,\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)} \text{ with } \sigma_{y\psi}^H = \sigma_{y\psi}^L - \gamma A_2(\tau^{**})\left(\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right). \end{aligned}$$

Above, $C_0(\tau)$ is given by

$$\begin{aligned} C_0(\tau) &= (1-\gamma)\,\overline{\rho}\,(\tau - A_1(\tau)) \\ &\quad + \frac{1}{2\phi^2}\left\{\tau + \frac{1-e^{-2\phi\tau}}{2\phi} - 2\frac{1-e^{-\phi\tau}}{\phi}\right\}\left(\gamma^2\sigma^2 - 2\gamma\sigma_{N,0}\sigma + \left(\sigma_{N,0}^2+\sigma_{N,1}^2\right)\right). \end{aligned}$$

Proof of Corollary 2. Let $\overline{p}_t = 1-\mathcal{N}\left(\underline{\psi};\widehat{\psi}_t + (1-\gamma)\,A_2(\tau^{**})\,(\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2)\,,\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)$. The claim follows from an application of Ito's Lemma, where we obtain

$$S_{M,t} = \frac{-\frac{\partial p_t}{\partial\widehat{\psi}_t}G_t^{no} + \left((1-\gamma)\,A_2(\tau^{**}) + \frac{1}{\overline{p}_t}\frac{\partial\overline{p}_t}{\partial\widehat{\psi}_t}\right)G_t^{yes}}{(1-p_t)\,G_t^{no} + p_t G_t^{yes}}. \tag{45}$$

Let also $p_{L,t}^N = \mathcal{N}\left(\underline{\psi};\widehat{\psi}_t+\sigma_{y\widehat{\psi}}^L,\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)$ and $p_{H,t}^N = 1-\mathcal{N}\left(\underline{\psi};\widehat{\psi}_t+\sigma_{y\widehat{\psi}}^H,\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)$, then

$$S_{M,t}^N = \frac{\left(A_2(\tau) + \frac{1}{p_{L,t}^N}\frac{\partial p_{L,t}^N}{\partial\widehat{\psi}}\right)K_t^{no} + \left((A_2(\tau) - \gamma A_2(\tau^{**})) + \frac{1}{p_{H,t}^N}\frac{\partial p_{H,t}^N}{\partial\widehat{\psi}}\right)K_t^{yes}}{(1-p_t)K_t^{no} + p_t K_t^{yes}}. \tag{46}$$

Proof of Proposition 6. First, we rewrite the M/B ratio of the old economy as

$$MB_t = \frac{G_t^{no} + p_t H_t}{\widetilde{G}_t^{no} + p_t\widetilde{H}_t},$$

where $H_t = G_t^{yes} - G_t^{no}$ and $\widetilde{H}_t = \widetilde{G}_t^{yes} - \widetilde{G}_t^{no}$. Taking the derivative $\partial MB_t/\partial p_t$, we find that M/B increases in p_t if and only if $H_t \widetilde{G}_t^{no} > G_t^{no} \widetilde{H}_t$. Substituting the closed-form expressions, we obtain the condition $h_{old} > 0$, where

$$\begin{aligned} h_{old} &= -\widetilde{\kappa} + A_2(\tau^{**})\widehat{\psi}_t + \frac{1}{2}(1-2\gamma) A_2(\tau^{**})^2 \widehat{\sigma}_t^2 && (47) \\ &\quad - \log\left(\frac{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2(\tau^{**})(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2), \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)}{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + (1-\gamma) A_2(\tau^{**})(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2), \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)}\right). && (48) \end{aligned}$$

We follow a similar derivation for the new economy's M/B ratio. First, we write

$$MB_t^N = \frac{K_t R_L^N + p_t \bar{J}_t}{\widetilde{G}_t^{no} + p_t \widetilde{H}_t},$$

where $\bar{J}_t = (1-\kappa)^{-\gamma} K_t^N R_H^N - K_t R_L^N$. Taking $\partial MB_t^N/\partial p_t$, we find that MB_t^N increases in p_t if and only if $\bar{J}_t \widetilde{G}_t^{no} - K_t \widetilde{H}_t R_L^N > 0$. Substituting, we obtain the condition $h_{new} > 0$, where

$$\begin{aligned} h_{new} &= -\gamma A_2(\tau^{**}) A_2(\tau) \widehat{\sigma}_t^2 - \log\left(\frac{\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + \sigma_{y\widehat{\psi}}^L, \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)}{\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)}\right) && (49) \\ &\quad - \log\left(\frac{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2(\tau^{**})(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2), \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)}{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2(\tau^{**})(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2) + \sigma_{y\psi}^L, \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)}\right). && (50) \end{aligned}$$

Proof of Proposition 7. First, we write $\frac{M^N}{B^N} = \frac{\Phi^N}{\widetilde{\pi}}$, where Φ^N and $\widetilde{\pi}$ are defined as the numerator and denominator in equation (25). Then,

$$\frac{\partial\left(\frac{M^N}{B^N}\right)}{\partial \widehat{\psi}_t} = \frac{\widetilde{\pi}\partial\Phi^N/\partial\widehat{\psi}_t - \Phi^N \partial\widetilde{\pi}/\partial\widehat{\psi}_t}{\widetilde{\pi}^2} > 0 \quad \text{if and only if} \quad S_{M,t}^N + S_{\pi,t} > 0,$$

where $S_{M,t}^N$ and $S_{\pi,t}$ are defined above. The probability of adoption as of time t^* is given by

$$p_{t^*} = \int_{f\left(\kappa,\gamma,\widehat{\sigma}_{t^*}^2;\tau^*\right)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx,$$

where

$$\begin{aligned} f\left(\kappa,\gamma,\widehat{\sigma}_{t^*}^2;\tau^*\right) &= -\log(1-\kappa)/A_2(\tau^{**}) \left(\frac{\left(\widehat{\sigma}_{t^*}^2\right)^{-1} + \left(\frac{\phi}{\sigma_{N,1}}\right)^2 (t^{**}-t^*)}{\widehat{\sigma}_{t^*}^2 \left(\frac{\phi}{\sigma_{N,1}}\right)^2 (t^{**}-t^*)}\right)^{\frac{1}{2}} \\ &\quad + \frac{\frac{1}{2}(\gamma-1) A_2(\tau^{**})}{\left(\frac{\phi}{\sigma_{N,1}}\right)(t^{**}-t^*)^{\frac{1}{2}} \left(1 + \widehat{\sigma}_{t^*}^2 \left(\frac{\phi}{\sigma_{N,1}}\right)^2 (t^{**}-t^*)\right)^{\frac{1}{2}}} \end{aligned}$$

Thus, $p_{t^*} \to 0$ if and only if $f\left(\kappa, \gamma, \widehat{\sigma}_{t^*}^2; \tau^*\right) \to \infty$. This happens when $\kappa \to 1$, $\gamma \to \infty$, $T \to \infty$, $t^{**} - t \to 0$, and, if $\kappa > 0$, when $\widehat{\sigma}_{t^*}^2 \to 0$. In all of these cases, the formulas for the various quantities in $S_{M,t}^N + S_{\pi,t}$ imply that this sum becomes positive.

Proof of Corollary 3. Immediate from Proposition 5 for $p_{t^{**}} = 1$ and $p_{t^{**}} = 0$. In Corollary 3, $\overline{C}_0(\tau^{**}) = C_0(\tau^{**}) - \overline{A}_0(\tau^{**})$.

Optimal Stopping Time.

Proposition 8: The value function in equation (35) is given by

$$\mathcal{V}\left(B_t, \rho_t, \widetilde{\psi}_t, \widehat{\sigma}_t^2, t; T\right) = B_t^{1-\gamma} e^{(1-\gamma)A_1(t)\rho_t} \mathcal{V}_2\left(\widehat{\psi}_t, t; T\right), \tag{51}$$

where $\mathcal{V}_2\left(\widehat{\psi}_t, t; T\right)$ satisfies the PDE

$$0 = \frac{\partial \mathcal{V}_2}{\partial t} + \left((1-\gamma) A_1(T-t)\phi\overline{\rho} + \frac{1}{2}(1-\gamma)^2 A_1(T-t)^2\sigma^2\right)\mathcal{V}_2 + \frac{1}{2}\frac{\partial^2 \mathcal{V}_2}{\partial \widehat{\psi}^2}\left(\widehat{\sigma}_t^2 \frac{\phi}{\sigma_{N,1}}\right)^2,$$

with the boundary conditions $\mathcal{V}_2\left(\widehat{\psi}_T, T\right) = \frac{1}{1-\gamma}$ if $t^{**} > T$ and

$$\mathcal{V}_2\left(\widehat{\psi}_t, t; T\right) \geq \frac{(1-\kappa)^{1-\gamma}}{1-\gamma} e^{A_0(\tau) + (1-\gamma)A_2(\tau)\widehat{\psi}_t + \frac{1}{2}(1-\gamma)^2 A_2(\tau)^2 \widehat{\sigma}_t^2},$$

where the equality holds at $t = t^{**}$.

Proof: Since $\widehat{\sigma}_t^2$ is a deterministic function of time, we write the value function simply as $\mathcal{V}(B_t, \rho_t, \widehat{\psi}_t, t; T)$. For $t \leq t^{**}$, $\mathcal{V}$ must satisfy the Bellman equation

$$0 = \frac{\partial \mathcal{V}}{\partial t} + \frac{\partial \mathcal{V}}{\partial B_t} E_t\left[dB_t\right] + \frac{\partial \mathcal{V}}{\partial \rho} E_t\left[d\rho\right] + \frac{\partial \mathcal{V}}{\partial \widetilde{\psi}} E_t\left[d\widetilde{\psi}\right] + \frac{1}{2}\frac{\partial^2 \mathcal{V}}{\partial \rho^2} E_t\left[d\rho_t^2\right] + \frac{1}{2}\frac{\partial^2 \mathcal{V}}{\partial \widehat{\psi}^2} E_t\left[d\widehat{\psi}_t^2\right] + \frac{\partial^2 \mathcal{V}}{\partial \rho \partial \widehat{\psi}} E_t\left[d\rho d\widehat{\psi}_t\right],$$

with the boundary conditions $\mathcal{V}(B_t, \rho_t, \widehat{\psi}_t, t; T) \geq V(B_t(1-\kappa), \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t, t; T)$ (and equality at t^{**}) and $\mathcal{V}(B_T, \rho_T, \widehat{\psi}_T, \widehat{\sigma}_T, T; T) = B_T^{1-\gamma}/(1-\gamma)$ if $T < t^{**}$. It is easy to verify that this Bellman equation is satisfied by the value function (51) with $\mathcal{V}_2$ satisfying the PDE and the boundary conditions given in Proposition 8.

Technical Appendix to Accompany Technological Revolutions and Stock Prices

Ľuboš Pástor
University of Chicago, *CEPR*, and *NBER*

Pietro Veronesi
University of Chicago, *CEPR*, and *NBER*

November 4, 2005

Technical Appendix

This appendix contains more detailed proofs than the ones sketched in the article.

Lemma 1: For later reference, we prove a more general version of Lemma 1. In particular, we cover three cases: (i) the new economy does not exist, and learning only occurs only by observing the old economy; (ii) the new economy exists, and learning occurs for $t \in [t^*, t^{**}]$; (iii) the new economy exists, adoption takes place at t^{**} and learning occurs for $t \geq t^{**}$. The learning dynamics for $t > t^{**}$ in the case of no adoption at t^{**} is identical to case (ii). For $t \geq t^*$ we then have

$$d\widehat{\psi}_t = \widehat{\sigma}_t^2 c \frac{\phi}{\sigma} d\widetilde{Z}_{0,t} + c_N \widehat{\sigma}_t^2 \frac{\phi}{\sigma_{N,1}} \left(1 - c\frac{\sigma_{N,0}}{\sigma}\right) d\widetilde{Z}_{1,t} \tag{B1}$$

$$\frac{d\widehat{\sigma}_t^2}{dt} = -\left(\widehat{\sigma}_t^2\right)^2 g \tag{B2}$$

where g, c and c_N are constants given by

$$g = \left(\left(\frac{c\phi}{\sigma}\right)^2 + c_N \left(\frac{\phi}{\sigma_{N,1}}\right)^2 \left(1 - c\frac{\sigma_{N,0}}{\sigma}\right)^2\right) \tag{B3}$$

$$(c, c_N) = \begin{cases} (1,0) & \text{if} \quad \text{only old economy exists} \\ (1,1) & \text{if} \quad t \geq t^{**} \text{ and adoption occurs at } t^{**} \\ (0,1) & \quad\quad \text{otherwise} \end{cases} \tag{B4}$$

This implies that

$$\widehat{\sigma}_t^2 = \begin{cases} \left(\widehat{\sigma}_{t^{**}}^{-2} + g\left(t - t^{**}\right)\right)^{-1} & \text{if} \quad t \geq t^{**} \text{ and switch occurs at } t^{**} \\ \left(\widehat{\sigma}_{t^{*}}^{-2} + g\left(t - t^{*}\right)\right)^{-1} & \quad\quad \text{otherwise} \end{cases} \tag{B5}$$

Proof: We consider only case (ii) and (iii). The simpler case (i) can be shown using similar steps. In these two cases, the new economy exists and thus the observation equations are

$$\begin{aligned} d\rho_t &= \phi\left(\overline{\rho} + c\psi - \rho_t\right) dt + \sigma_0 dZ_{0,t} \\ d\rho_t^N &= \phi\left(\overline{\rho} + \psi - \rho_t^N\right) dt + \sigma_{N,0} dZ_{0,t} + \sigma_{N,1} dZ_{1,t} \end{aligned}$$

where c is given in (B4). Defining $\mathbf{s}_t = (\rho_t, \rho_t^N)'$, this can be written compactly as

$$d\mathbf{s}_t = (\mathbf{A} + \mathbf{B}\mathbf{z} + \mathbf{C}\psi)\, dt + \mathbf{\Sigma} d\mathbf{Z}$$

where $\mathbf{C} = (c\phi, \phi)'$ and

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma & 0 \\ \sigma_{N,0} & \sigma_{N,1} \end{pmatrix}$$

Liptser and Shiryaev (1977) show that the process for $\widehat{\psi}_t = E_t[\psi]$ is given by

$$d\widehat{\psi}_t = \widehat{\sigma}_t^2 \mathbf{C}' \left(\mathbf{\Sigma}'\right)^{-1} d\widetilde{\mathbf{Z}} \tag{B6}$$

where $\widetilde{\mathbf{Z}}_t = (\widetilde{Z}_{0,t}, \widetilde{Z}_{1,t})'$ follows the process

$$d\widetilde{\mathbf{Z}}_t = \mathbf{\Sigma}^{-1} \left(\begin{array}{c} d\rho_t \\ d\rho_t^N \end{array} - E_t \left[\begin{array}{c} d\rho_t \\ d\rho_t^N \end{array} \right] \right)$$

and

$$\frac{d\widehat{\sigma}_t^2}{dt} = -\left(\widehat{\sigma}_t^2\right)^2 \mathbf{C}' \left(\mathbf{\Sigma}\mathbf{\Sigma}'\right)^{-1} \mathbf{C}$$

Substituting $\mathbf{C}$ and $\mathbf{\Sigma}$, we find immediately

$$\mathbf{C}' \left(\mathbf{\Sigma}'\right)^{-1} = \left(c\frac{\phi}{\sigma}, -c\phi\frac{\sigma_{N,0}}{\sigma\sigma_{N,1}} + \frac{\phi}{\sigma_{N,1}}\right)$$

Substituting this expression in (B6) and defining $g = \mathbf{C}' \left(\mathbf{\Sigma}\mathbf{\Sigma}'\right)^{-1} \mathbf{C}$ we obtain (B1) and (B2) for $c_N = 1$. It is simple to verify that (B5) satisfies (B2), yielding the conclusion. Q.E.D.

It is convenient to rewrite the original processes under the filtered measure. Let $b_t = \log(B_t)$ and $b_t^N = \log\left(B_t^N\right)$. For $t > t^*$ we have

$$\begin{aligned}
db_t &= \rho_t dt && \text{(B7)}\\
d\rho_t &= \phi\left(\overline{\rho} + c\widehat{\psi}_t - \rho_t\right)dt + \sigma d\widetilde{Z}_{0,t} && \text{(B8)}\\
d\widehat{\psi} &= \widehat{\sigma}_t^2 c\frac{\phi}{\sigma} d\widetilde{Z}_{0,t} + c_N \widehat{\sigma}_t^2 \frac{\phi}{\sigma_{N,1}}\left(1 - c\frac{\sigma_{N,0}}{\sigma}\right) d\widetilde{Z}_{1,t} && \text{(B9)}\\
d\widehat{\sigma}_t^2 &= -\left(\widehat{\sigma}_t^2\right)^2 \left(\left(\frac{c\phi}{\sigma}\right)^2 + c_N\left(\frac{\phi}{\sigma_{N,1}}\right)^2\left(1 - c\frac{\sigma_{N,0}}{\sigma}\right)^2\right) dt && \text{(B10)}\\
db_t^N &= \rho_t^N dt && \text{(B11)}\\
d\rho_t^N &= \phi\left(\overline{\rho} + \widehat{\psi}_t - \rho_t^N\right)dt + \sigma_{N,0} d\widetilde{Z}_{0,t} + \sigma_{N,1} d\widetilde{Z}_{1,t} && \text{(B12)}
\end{aligned}$$

Lemma A1: Let $\tau = T - t$. The expectation in equation (6) is given by

$$V\left(B_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2, \tau\right) = E_t\left[\frac{B_T^{1-\gamma}}{1-\gamma}\right] = \frac{B_t^{1-\gamma}}{1-\gamma} e^{A_0(\tau) + (1-\gamma)A_1(\tau)\rho_t + (1-\gamma)A_2(\tau)\widehat{\psi}_t + \frac{1}{2}(1-\gamma)^2 A_2(\tau)^2 \widehat{\sigma}_t^2} \quad \text{(B13)}$$

where

$$\begin{aligned}
A_0(\tau) &= (1-\gamma)\overline{\rho}(\tau - A_1(\tau)) + \frac{\sigma^2}{2}\frac{(1-\gamma)^2}{\phi^2}\left\{\tau + \frac{1 - e^{-2\phi\tau}}{2\phi} - 2\frac{1 - e^{-\phi\tau}}{\phi}\right\}\\
A_1(\tau) &= \frac{1 - e^{-\phi\tau}}{\phi} \text{ and } A_2(\tau) = \tau - A_1(\tau)
\end{aligned}$$

Proof: By definition

$$V\left(b_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2, t; T\right) = (1-\gamma)^{-1} E_t\left[e^{(1-\gamma)b_T}\right]$$

Denoting $\mathbf{x}_t = \left(b_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2\right)$, the Feynman-Kac theorem shows that V has to satisfy the PDE

$$0 = \frac{\partial V}{\partial t} + \sum_i \frac{\partial V}{\partial x_i} E_t[dx_i] + \frac{1}{2}\sum_i\sum_j \frac{\partial^2 V}{\partial x_i \partial x_j} E_t[dx_i dx_j]$$

with boundary condition $V(\mathbf{x}_T) = (1-\gamma)^{-1} e^{(1-\gamma)x_{1,T}}$. Using (B7) - (B10) with $c = 1$ and $c_N = 0$, it is simple to verify that (B13) satisfies this PDE with the boundary condition. Finally, $A_2(\tau) > 0$

is immediate. Rewrite $A_2(\tau) = f(\tau) = \tau - \frac{1-e^{-\phi\tau}}{\phi}$. Note that $f(0) = 0$. Since $f'(\tau) = 1 - e^{-\phi\tau} > 0$, we have $f(\tau) > 0$ for every $\tau > 0$. QED.

Proof of Proposition 1: Since $\gamma > 1$ we have that V in (B13) is decreasing in $\widehat{\sigma}_t^2$. It immediately follows that $V\left(B_{t^*}(1-\kappa), \rho_{t^*}, 0, \widehat{\sigma}_{t^*}^2, \tau^*\right) < V(B_{t^*}, \rho_{t^*}, 0, 0, \tau^*)$. Q.E.D.

Proof of Proposition 2: Using (B13) it is immediate to verify that equation (13) follows from equation (14). Q.E.D.

To prove Proposition 3 we need the following lemmas, obtaining the closed form solution for the value function in equation (15) in the paper:

Lemma A2: The density of $\widehat{\psi}_{t^{**}}$ conditional on $\widehat{\psi}_t$ is normal and explicitly given by

$$\widehat{\psi}_{t^{**}}|_{\widehat{\psi}_t} \sim N\left(\widehat{\psi}_t, \sigma_{\widehat{\psi},t}^2\right)$$

where

$$\sigma_{\widehat{\psi},t}^2 = \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2$$

and $\widehat{\sigma}_t^2$ is given in (B5) for the case $t < t^{**}$.

Proof: The process for the posterior mean $\widehat{\psi}_t$ is a linear diffusion with deterministic volatility, as given in (B1). The integral representation is

$$\widehat{\psi}_{t^{**}} = \widehat{\psi}_t + \frac{\phi}{\sigma_{N,1}} \int_t^{t^{**}} \widehat{\sigma}_s^2 d\widetilde{Z}_{1,s}$$

which immediately implies that

$$\widehat{\psi}_{t^{**}}|\widehat{\psi}_t \sim N\left(\widehat{\psi}_t, \sigma_{\widehat{\psi},t}^2\right)$$

where

$$\sigma_{\widehat{\psi},t}^2 = \left(\frac{\phi}{\sigma_{N,1}}\right)^2 \int_t^{t^{**}} \left(\widehat{\sigma}_s^2\right)^2 ds$$

Using (B5) for $t < t^{**}$ we can compute

$$\int_t^{t^{**}} \left(\widehat{\sigma}_s^2\right)^2 ds = \frac{1}{(\phi/\sigma_{N,1})^2}\left[\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right]$$

Thus $\sigma_{\widehat{\psi},t}^2 = \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2$. In addition, it is then immediate that the probability of adoption is given by

$$p_t \equiv p\left(\widehat{\psi}_t, t\right) = \Pr\left(\widehat{\psi}_{t^{**}} > \underline{\psi}|\widehat{\psi}_t\right) = 1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \sigma_{\widehat{\psi},t}^2\right)$$

where $\mathcal{N}(.; a, s^2)$ the cumulative density function of a normal distribution with mean a and variance s^2. Q.E.D.

Lemma A3: The value function

$$\mathcal{V}_t = E_t\left[\max_{yes,no} E_{t^{**}}\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right]\right]$$

at time $t^* \leq t < t^{**}$ is given by

$$\mathcal{V}\left(B_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2; \tau\right) = \frac{B_t^{1-\gamma}}{1-\gamma}\left\{(1-p_t)\, G_t^{no} + p_t G_t^{yes}\right\} \tag{B14}$$

where

$$\begin{aligned} G_t^{no} &= e^{A_0(\tau)+(1-\gamma)A_1(\tau)\rho_t} \\ G_t^{yes} &= G_t^{no}\,(1-\kappa)^{1-\gamma}\, R_t e^{(1-\gamma)A_2(\tau^{**})\widehat{\psi}_t + \frac{1}{2}(1-\gamma)^2 A_2(\tau^{**})^2 \widehat{\sigma}_t^2} \end{aligned}$$

and

$$R_t = \frac{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + (1-\gamma)\, A_2(\tau^{**})\, \sigma^2_{\widehat{\psi},t}, \sigma^2_{\widehat{\psi},t}\right)}{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \sigma^2_{\widehat{\psi},t}\right)} < 1 \tag{B15}$$

Proof: The value function is

$$\mathcal{V}_t = E_t\left[\max_{yes,no} E_{t^{**}}\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right]\right] = (1-p_t)\, E_t\left[\frac{W_T^{1-\gamma}}{1-\gamma}|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] + p_t E_t\left[\frac{W_T^{1-\gamma}}{1-\gamma}|\widehat{\psi}_{t^{**}} \geq \underline{\psi}\right]$$

as the adoption at t^{**} occurs if and only if $\widehat{\psi}_{t^{**}} \geq \underline{\psi}$. Starting with the first expectation, we can use the law of iterated expectations

$$E_t\left[\frac{W_T^{1-\gamma}}{1-\gamma}|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] = E_t\left[E_{t^{**}}\left[\frac{W_T^{1-\gamma}}{1-\gamma}|\widehat{\psi}_{t^{**}} < \underline{\psi}\right]|\widehat{\psi}_{t^{**}} < \underline{\psi}\right]$$

We can use again equation (B13) to compute the inner expectation. In fact, if $\widehat{\psi}_{t^{**}} < \underline{\psi}$ the technology does not change at t^{**}. Moreover, eqn (B8) - (B9) show that ρ_t and $\widehat{\psi}_t$ are independent as $c = 0$ (see eqn. B4). Thus, Lemma A2 implies

$$E_{t^{**}}\left[\frac{W_T^{1-\gamma}}{1-\gamma}|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] = V\left(B_{t^{**}}, \rho_{t^{**}}, 0, 0, t^{**}; T\right) = \frac{B_{t^{**}}^{1-\gamma}}{1-\gamma} e^{A_0(t^{**};T)+A_1(t^{**};T)\rho_{t^{**}}}$$

Thus,

$$\begin{aligned} E_t\left[\frac{B_{t^{**}}^{1-\gamma}}{1-\gamma} e^{A_0(t^{**};T)+A_1(t^{**};T)\rho_{t^{**}}}|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] &= E_t\left[\frac{B_{t^{**}}^{1-\gamma}}{1-\gamma} e^{A_0(t^{**};T)+A_1(t^{**};T)\rho_{t^{**}}}\right] \\ &= \frac{B_t^{1-\gamma}}{1-\gamma} e^{A_0(t;T)+A_1(t;T)\rho_t} \end{aligned}$$

where the first equality stems from the independence of ρ_t and $\widehat{\psi}_t$, and the second equality stems from an application of Feynman - Kac thorem, similar to the argument used in Lemma A1.

The second expectation is more involved, as until t^{**} capital employs the old technology, and only then it switches to the new technology. In addition, the switch occurs only if $\widehat{\psi}_{t^{**}}$ is high enough, and this must be taken into account in the computation. Using again the law of iterated expectations, we have

$$\begin{aligned} E_t\left[\frac{W_T^{1-\gamma}}{1-\gamma}|\widehat{\psi}_{t^{**}} \geq \underline{\psi}\right] &= E_t\left[E_{t^{**}}\left[\frac{W_T^{1-\gamma}}{1-\gamma}|\widehat{\psi}_{t^{**}} > \underline{\psi}\right]|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] \\ &= E_t\left[V\left(B_{t^{**}}(1-\kappa), \rho_{t^{**}}, \widehat{\psi}_{t^{**}}, \widehat{\sigma}^2_{t^{**}}, t^{**}; T\right)|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] \end{aligned}$$

where the second equality stems from Lemma A1 and the fact that if $\widehat{\psi}_{t^{**}} > \underline{\psi}$, the adoption occurs. We can use the explicit formula for $V(.)$ to compute this expectation. In particular, from (B7) - (B10), $\widehat{\psi}_t$ is independent of both ρ_t and b_t, and $\widehat{\sigma}^2_{t^{**}}$ is a known constant. Thus, we can write

$$\begin{aligned} & E_t\left[V\left(B_{t^{**}}(1-\kappa), \rho_{t^{**}}, \widehat{\psi}_{t^{**}}, \widehat{\sigma}^2_{t^{**}}, t^{**}; T\right)|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] \\ = & \frac{(1-\kappa)^{1-\gamma}}{1-\gamma} E_t\left[e^{(1-\gamma)b_{t**}+A_0(t^{**};T)+(1-\gamma)A_1(t^{**};T)\rho_{t**}+\frac{1}{2}(1-\gamma)^2 A_2(t^{**};T)^2\widehat{\sigma}^2_{t^{**}}}\right] \\ & \times E_t\left[e^{(1-\gamma)A_2(t^{**};T)\widehat{\psi}_{t**}}|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] \\ = & e^{(1-\gamma)b_t+A_0(t;T)+(1-\gamma)A_1(t;T)\rho_t+\frac{1}{2}(1-\gamma)^2 A_2(t^{**};T)^2\widehat{\sigma}^2_{t^{**}}} E_t\left[e^{(1-\gamma)A_2(t^{**};T)\widehat{\psi}_{t**}}|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] \end{aligned}$$

Since from Lemma A2, $\widehat{\psi}_{t^{**}} \sim N\left(\widehat{\psi}_t, \widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right)$ we have that the conditional density required to compute the last expectation is given by

$$f\left(\widehat{\psi}_{t**}|\widehat{\psi}_{t**} > \underline{\psi}\right) = \frac{f\left(\widehat{\psi}_{t**}; \widehat{\psi}_t, \widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right) 1_{\{\widehat{\psi}_{t**} > \underline{\psi}\}}}{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right)}$$

Using this density, we find

$$\begin{aligned} E\left[e^{(1-\gamma)A_2(t^{**};T)\widehat{\psi}_{t**}}|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] &= \frac{1}{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right)} \int_{\underline{\psi}}^{\infty} e^{(1-\gamma)A_2(t^{**};T)\widehat{\psi}_{t**}} f\left(\widehat{\psi}_{t**}\right) d\widehat{\psi}_{t^{**}} \\ &= e^{\frac{1}{2}(1-\gamma)^2 A_2^2(\tau^{**})\left(\widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right)+(1-\gamma)A_2(\tau^{**})\widehat{\psi}_t} R\left(\widehat{\psi}_t\right) \end{aligned}$$

where $R\left(\widehat{\psi}_t\right) = R_t$ is given in (B15). Putting all these elements together, we obtain (B14).

Lemma A4: $G_t^{yes} < G_t^{no}$.

Proof: Consider the expression

$$J_t = E_t\left[e^{(1-\gamma)\log(1-\kappa)+(1-\gamma)A_2(t^{**};T)\widehat{\psi}_{t**}+\frac{1}{2}(1-\gamma)^2 A_2(t^{**};T)^2\widehat{\sigma}^2_{t^{**}}}|\widehat{\psi}_{t^{**}} > \underline{\psi}\right]$$

Using the definition of $\underline{\psi}$ in equation (13) of the paper, this can be written as

$$\begin{aligned} J_t &= E_t\left[e^{-(1-\gamma)A_2(\tau^{**})\left[-\frac{\log(1-\kappa)}{A_2(\tau^*)}-\widehat{\psi}_{t**}-\frac{1}{2}(1-\gamma)A_2(t^{**};T)\widehat{\sigma}^2_{t^{**}}\right]}|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] \\ &= E_t\left[e^{(1-\gamma)A_2(\tau^{**})\left[\widehat{\psi}_{t**}-\underline{\psi}\right]}|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] \end{aligned}$$

Thus, $J_t < 1$, as it is the expectation of a random variable that is constrained to be less than 1. By using the same steps as in Lemma A3, we find

$$\begin{aligned} J_t &= E_t\left[e^{(1-\gamma)A_2(\tau^{**})\left[\widehat{\psi}_{t^{**}}-\underline{\psi}\right]}|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] \\ &= e^{-(1-\gamma)A_2(\tau^{**})\underline{\psi}}E_t\left[e^{(1-\gamma)A_2(\tau^{**})\widehat{\psi}_{t^{**}}}|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] \\ &= e^{-(1-\gamma)A_2(\tau^{**})\underline{\psi}+(1-\gamma)A_2(\tau^*)\widehat{\psi}_t+\frac{1}{2}(1-\gamma)^2A_2(\tau^*)^2\left(\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)} \times R_t \\ &= e^{(1-\gamma)\log(1-\kappa)+(1-\gamma)A_2(\tau^*)\widehat{\psi}_t+\frac{1}{2}(1-\gamma)^2A_2(\tau^*)^2\widehat{\sigma}_t^2} \times R_t \\ &= \frac{G_t^{yes}}{G_t^{no}} \end{aligned}$$

yielding the conclusion. Q.E.D.

Proposition 3: Experimenting is always optimal at time t^*, that is

$$\mathcal{V}\left(B_{t^*}, \rho_{t^*}, 0, \widehat{\sigma}_{t^*}^2; \tau^*\right) > V\left(B_{t^*}, \rho_{t^*}, 0, 0; \tau^*\right)$$

where $V\left(B_{t^*}, \rho_{t^*}, 0, 0; \tau^*\right)$ is defined in equation (B13).

Proof: Since $G_t^{yes} < G_t^{no}$, the result follows from the fact that we can rewrite $V\left(B_{t^*}, \rho_{t^*}, 0, 0; \tau^*\right) = \frac{B_{t^*}^{1-\gamma}}{1-\gamma}G_{t^*}^{no}$ and $\gamma > 1$. Q.E.D.

Proof of Proposition 4: The proof is identical to the one of Lemma A3, where "$(1-\gamma)$" is substituted with "$-\gamma$". Using this fact, we have

$$\pi_t = \lambda^{-1}B_t^{-\gamma}\left\{(1-p_t)\,\widetilde{G}_t^{no} + p_t\widetilde{G}_t^{yes}\right\} \tag{B16}$$

where

$$\widetilde{G}_t^{no} = e^{\overline{A}_0(\tau)-\gamma A_1(\tau)\rho_t} \tag{B17}$$

$$\widetilde{G}_t^{yes} = \widetilde{G}_t^{no}\,(1-\kappa)^{-\gamma}\,\widetilde{R}_t e^{-\gamma A_2(\tau^{**})\widehat{\psi}_t+\frac{1}{2}\gamma^2A_2(\tau^{**})^2\widehat{\sigma}_t^2} \tag{B18}$$

and

$$\widetilde{R}_t = \frac{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2\left(\tau^{**}\right)\sigma_{\widehat{\psi},t}^2, \sigma_{\widehat{\psi},t}^2\right)}{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \sigma_{\widehat{\psi},t}^2\right)} < 1 \tag{B19}$$

In this proposition,

$$\overline{A}_0\left(\tau\right) = -\gamma\overline{\rho}\left(\tau - A_1\left(\tau\right)\right) + \frac{\sigma^2}{2}\frac{\gamma^2}{\phi^2}\left\{\tau + \frac{1-e^{-2\phi\tau}}{2\phi} - 2\frac{1-e^{-\phi\tau}}{\phi}\right\}$$

Q.E.D.

Proof of Corollary 1: The corollary follows from an application of Ito's Lemma, so that

$$\frac{d\pi_t}{\pi_t} = -\boldsymbol{\sigma}_{\pi,t}d\widetilde{\mathbf{Z}}_t$$

where

$$\boldsymbol{\sigma}_{\pi,t} = \gamma A_1(\tau)\,\boldsymbol{\sigma} + S_{\pi,t}\widetilde{\boldsymbol{\sigma}}_{\psi,t}$$

and

$$S_{\pi,t} = \frac{\left(\gamma A_2(\tau^{**}) - \frac{1}{\widetilde{p}}\frac{\partial \widetilde{p}}{\partial \widehat{\psi}}\right)\widetilde{G}_t^{yes} + \frac{\partial p}{\partial \widehat{\psi}}\widetilde{G}_t^{no}}{(1-p_t)\,\widetilde{G}_t^{no} + p_t\widetilde{G}_t^{yes}} \tag{B20}$$

where

$$\widetilde{p}_t \equiv \widetilde{p}\left(\widehat{\psi}_t, t\right) = 1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2(\tau^{**})\,\sigma^2_{\widehat{\psi},t}, \sigma^2_{\widehat{\psi},t}\right)$$

and $\boldsymbol{\sigma} = (\sigma, 0)$, $\widetilde{\boldsymbol{\sigma}}_{\psi} = \left(0, \widehat{\sigma}_t^2 \frac{\phi}{\sigma_{N,1}}\right)$. Q.E.D.

Proof of Proposition 5 (old economy): The result about the old economy is immediate from the pricing formula $M_t = E_t[\pi_T B_T]/\pi_t = E_t\left[B_T^{1-\gamma}\right]/\pi_t$, and the results in Lemma A3 and Proposition 4. Q.E.D.

For better referencing, it is convenient to restate Proposition 5 for the new economy:

Proposition 5 (new economy) Let $\tau = T - t$. For $t^* \le t < t^{**}$, the market to book ratio of the new economy is given by

$$\frac{M_t^N}{B_t^N} = \frac{(1-p_t)\,K^{no} + p_t K^{yes}}{(1-p_t)\,\widetilde{G}_t^{no} + p_t\widetilde{G}_t^{yes}} \tag{B21}$$

where $\widetilde{G}_t^{no}$ and $\widetilde{G}_t^{yes}$ are given in Proposition 3, and

$$\begin{aligned} K^{no} &= K_t R_{L,t}^N \\ K^{yes} &= (1-\kappa)^{-\gamma} K_t^N R_{H,t}^N \end{aligned}$$

$$\begin{aligned} K_t &= e^{C_0(\tau) - \gamma A_1(\tau)\rho_t + A_1(\tau)\rho_t^N + A_2(\tau)\widehat{\psi}_t + \frac{1}{2}A_2^2(\tau)\widehat{\sigma}_t^2} \\ K_t^N &= K_t e^{-\gamma A_2(\tau^{**})\widehat{\psi}_t + \frac{1}{2}\gamma A_2(\tau^{**})(\gamma A_2(\tau^{**}) - 2A_2(\tau))\widehat{\sigma}_t^2} \end{aligned}$$

and

$$\begin{aligned} R_{L,t}^N &= \frac{\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + \sigma^L_{y\widehat{\psi}}, \sigma^2_{\widehat{\psi},t}\right)}{\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \sigma^2_{\widehat{\psi},t}\right)} \text{ with } \sigma^L_{y\widehat{\psi}} = A_2(\tau)\,\widehat{\sigma}_t^2 - A_2(\tau^{**})\,\widehat{\sigma}^2_{t^{**}} \\ R_{H,t}^N &= \frac{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + \sigma^H_{y\widehat{\psi}}, \sigma^2_{\widehat{\psi},t}\right)}{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \sigma^2_{\widehat{\psi},t}\right)} \text{ with } \sigma^H_{y\psi} = \sigma^L_{y\psi} - \gamma A_2(\tau^{**})\,\sigma^2_{\widehat{\psi},t} \end{aligned}$$

Above, $C_0(\tau)$ is given by

$$\begin{aligned} C_0(\tau) &= (1-\gamma)\,\overline{\rho}\,(\tau - A_1(\tau)) \\ &\quad + \frac{1}{2\phi^2}\left\{\tau + \frac{1-e^{-2\phi\tau}}{2\phi} - 2\frac{1-e^{-\phi\tau}}{\phi}\right\}\left(\gamma^2\sigma^2 - 2\gamma\sigma_{N,0}\sigma + \left(\sigma^2_{N,0} + \sigma^2_{N,1}\right)\right) \end{aligned}$$

We start the proof with two lemmas:

Lemma A5: For $t \geq t^{**}$, let $\tau = T - t$. Then

$$V^N\left(b_t, b_t^N, \rho_t, \rho_t^N, \widehat{\psi}_t, \widehat{\sigma}_t^2, \tau\right) \equiv E_t\left[e^{-\gamma b_T + b_T^N}\right]$$

is given by

$$V^N\left(b_t, b_t^N, \rho_t, \rho_t^N, \widehat{\psi}_t, \widehat{\sigma}_t^2, \tau\right) = e^{-\gamma b_t + b_t^N + C_0(\tau) - \gamma A_1(\tau)\rho_t + A_1(\tau)\rho_t^N + A_2(\tau)\widehat{\psi}_t + \frac{1}{2}(1-c\gamma)^2 A_2^2(\tau)\widehat{\sigma}_t^2} \tag{B22}$$

where $c = 1$ if the adoption occurred at time t^{**} , and 0 otherwise, $A_1(.)$ and $A_2(.)$ are as in Lemma A1, and

$$C_0(\tau) = (1-\gamma)\,\overline{\rho}\,(\tau - A_1(\tau)) + \frac{1}{\phi^2}\left(\tau + \frac{1-e^{-2\phi\tau}}{2\phi} - 2A_1(\tau)\right)\frac{1}{2}(\sigma^*)^2$$

and

$$(\sigma^*)^2 = \gamma^2\sigma^2 + \sigma_{N,0}^2 + \sigma_{N,1}^2 - 2\gamma\sigma_{N,0}\sigma$$

Proof: As in Lemma A1, denoting $\mathbf{x}_t = \left(b_t, b_t^N, \rho_t, \rho_t^N, \widehat{\psi}_t, \widehat{\sigma}_t^2\right)$, the Feynman-Kac theorem shows that V^N has to satisfy the PDE

$$0 = \frac{\partial V^N}{\partial t} + \sum_i \frac{\partial V^N}{\partial x_i} E_t[dx_i] + \frac{1}{2}\sum_i\sum_j \frac{\partial^2 V^N}{\partial x_i \partial x_j} E_t[dx_i dx_j]$$

with the boundary condition $V(\mathbf{x}_T) = (1-\gamma)^{-1} e^{-\gamma x_{1,T} + x_{2,T}}$. Using (B7) - (B12) for the cases where $c = 1$ or $c = 0$ (with $c_N = 1$) in Lemma 1, it is simple to verify that (B22) satisfies this PDE with the boundary condition provided. Q.E.D.

Lemma A6: Define

$$y_{t^{**}} = -\gamma b_{t^{**}} + b_{t^{**}}^N - \gamma A_1(\tau^{**})\,\rho_{t^{**}} + A_1(\tau^{**})\,\rho_{t^{**}}^N + (1-c_1)\,A_2(\tau^{**})\,\widehat{\psi}_{t^{**}}$$

where $c_1 > 0$ is a constant. Then

$$\begin{pmatrix} y_{t^{**}} \\ \widehat{\psi}_{t^{**}} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{y,t} \\ \widehat{\psi}_t \end{pmatrix}, \begin{pmatrix} \sigma_y^2 & \sigma_{y\psi} \\ \sigma_{y\psi} & \sigma_{\widehat{\psi},t}^2 \end{pmatrix}\right)$$

where

$$\begin{aligned}
\mu_{y,t} &= -\gamma b_t + b_t^N + (1-\gamma)\,\overline{\rho} a(t) - \gamma A_1(\tau)\,\rho_t + A_1(\tau)\,\rho_t^N + (A_2(\tau) - c_1 A_2(\tau^{**}))\,\widehat{\psi}_t \\
\sigma_y^2 &= (1-c_1)^2 A_2(\tau^{**})^2\left(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right) + a(t)^2\,\widehat{\sigma}_t^2 + 2A_2(\tau^{**})(1-c_1)\,a(t)\,\widehat{\sigma}_t^2 + (\sigma^*)^2 a_2(t) \\
\sigma_{\widehat{\psi},t}^2 &= \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2 \\
\sigma_{y\psi} &= (1-c_1)\,A_2(\tau^{**})\left(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right) + a(t)\,\widehat{\sigma}_t^2
\end{aligned}$$

and

$$
\begin{aligned}
a(t) &= t^{**} - t - \frac{e^{-\phi(T-t^{**})} - e^{-\phi(T-t)}}{\phi} && \text{(B23)} \\
a_2(t) &= \frac{1}{\phi^2}\left(t^{**} - t + \frac{e^{-2\phi(T-t^{**})} - e^{-2\phi(T-t)}}{2\phi} - 2\frac{e^{-\phi(T-t^{**})} - e^{-\phi(T-t)}}{\phi}\right) && \text{(B24)} \\
(\sigma^*)^2 &= \gamma^2\sigma^2 + \sigma_{N,0}^2 + \sigma_{N,1}^2 - 2\gamma\sigma_{N,0}\sigma && \text{(B25)}
\end{aligned}
$$

Proof: The proof of this lemma is rather lengthy, and so it is provided separately below.

Proof of Proposition 5 (new economy): The pricing formula is $M_t^N = E_t\left[\pi_T B_T^N\right]/\pi_t$. Thus, we need to compute

$$
E_t\left[B_T^{-\gamma}B_T^N\right] = (1-p_t)\,E_t\left[B_T^{-\gamma}B_T^N|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] + p_t E_t\left[B_T^{-\gamma}B_T^N|\widehat{\psi}_{t^{**}} > \underline{\psi}\right] \qquad \text{(B26)}
$$

Starting with the first expectation, note that if $\widehat{\psi}_{t^{**}} < \underline{\psi}$, no adoption occurs at t^{**}. Thus,

$$
\begin{aligned}
E_t\left[B_T^{-\gamma}B_T^N|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] &= E_t\left[E_{t^{**}}\left[B_T^{-\gamma}B_T^N|\widehat{\psi}_{t^{**}} < \underline{\psi}\right]|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] \\
&= E_t\left[V^N\left(b_{t^{**}}, b_{t^{**}}^N, \rho_{t^{**}}, \rho_{t^{**}}^N, \widehat{\psi}_{t^{**}}, \widehat{\sigma}_{t^{**}}^2, t^{**}; T\right)|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] \\
&= e^{C_0(\tau^{**}) + \frac{1}{2}A_2^2(\tau^{**})\widehat{\sigma}_{t^{**}}^2} \\
&\quad \times E_t\left[e^{-\gamma b_{t^{**}} + b_{t^{**}}^N - \gamma A_1(\tau^{**})\rho_{t^{**}} + A_1(\tau^{**})\rho_{t^{**}}^N + A_2(\tau^{**})\widehat{\psi}_{t^{**}}}|\widehat{\psi}_{t^{**}} < \underline{\psi}\right]
\end{aligned}
$$

where the first equality stems from the law of iterated expectations, the second from the fact that $\widehat{\psi}_{t^{**}}$ is known at t^{**}, the third from Lemma A5, with $c = 0$ as the adoption does not occur at t^{**}. Note that the exponent in the expectation is simply $y_{t^{**}}$ in Lemma A6 with $c_1 = 0$. For notational convenience, let

$$
a_0(t) = (1-\gamma)\,\overline{\rho}a(t).
$$

Using Lemma A6 with $c_1 = 0$ and denoting by L the corresponding quantities in Lemma A6 for this case, we can compute

$$
E\left[e^{y_{t^{**}}}|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] = \frac{\int_{-\infty}^{\underline{\psi}} E\left[e^{y_{t^{**}}}|\widehat{\psi}_{t^{**}}\right] f\left(\widehat{\psi}_{t^{**}}; \widehat{\psi}_t, \sigma_{\widehat{\psi},t}^2\right) d\widehat{\psi}_{t^{**}}}{\Pr\left(\widehat{\psi}_{t^{**}} < \underline{\psi}\right)}
$$

where $f\left(\widehat{\psi}_{t^{**}}; \widehat{\psi}_t, \sigma_{\widehat{\psi},t}^2\right)$ is the density of a normal with mean $\widehat{\psi}_t$ and variance $\sigma_{\widehat{\psi},t}^2$. The rules of the conditional normal distribution yield the following expression for this expectation:

$$
E_t\left[e^{y_{t^{**}}}|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] = B_t^{-\gamma}B_t^N e^{a_0(t) - \gamma A_1(t;T)\rho_t + A_1(t;T)\rho_t^N + A_2(t;T)\widehat{\psi}_t + \frac{1}{2}\sigma_{Ly}^2} R_{L,t}^N
$$

where $R_{L,t}^N$ is given in Proposition 5. So, finally, the first expectation is given by

$$
\begin{aligned}
E_t\left[B_T^{-\gamma}B_T^N|\widehat{\psi}_{t^{**}} < \underline{\psi}\right] &= B_t^{-\gamma}B_t^N e^{C_0(t^{**};T) + \frac{1}{2}A_2^2(t^{**};T)\widehat{\sigma}_{t^{**}}^2} e^{a_0(t) - \gamma A_1(t;T)\rho_t + A_1(t;T)\rho_t^N + A_2(t;T)\widehat{\psi}_t + \frac{1}{2}\sigma_{L,y}^2} R_{L,t}^N \\
&= B_t^{-\gamma}B_t^N e^{C_0(t;T) + \frac{1}{2}A_2^2(t;T)\widehat{\sigma}_t^2 - \gamma A_1(t;T)\rho_t + A_1(t;T)\rho_t^N + A_2(t;T)\widehat{\psi}_t} R_{L,t}^N
\end{aligned}
$$

where the second equality is obtained from the first after some tedious algebra.

We now turn to the second expectation in (B26). The methodology is the same as before, although now we must set $c = 1$ in (B22) and note that $B_{t^{**}} = (1-\kappa) B_{t^{**}_{-}}$, which implies $b_{t^{**}} = b_{t^{**}_{-}} + \log(1-\kappa)$. Specifically, we have that for $t \le t^{**}$

$$\begin{aligned} E_t\left[B_T^{-\gamma} B_T^N | \widehat{\psi}_{t^{**}} \ge \underline{\psi}\right] &= E_t\left[V^N\left(\log(1-\kappa) + b_{t^{**}_{-}}, b_{t^{**}}^N, \rho_{t^{**}}, \rho_{t^{**}}^N, \widehat{\psi}_{t^{**}}, \widehat{\sigma}_{t^{**}}^2, \tau\right) | \widehat{\psi}_{t^{**}} \ge \underline{\psi}\right] \\ &= (1-\kappa)^{-\gamma} e^{C_0(\tau^{**}) + \frac{1}{2}(1-\gamma)^2 A_2^2(\tau^{**})\widehat{\sigma}_{t^{**}}^2} \times \\ &\quad \times E_t\left[e^{-\gamma b_{t^{**}} + b_{t^{**}}^N - \gamma A_1(\tau^{**})\rho_{t^{**}} + A_1(\tau^{**})\rho_{t^{**}}^N + (1-\gamma)A_2(\tau^{**})\widehat{\psi}_{t^{**}}} | \widehat{\psi}_{t^{**}} \ge \underline{\psi}\right] \end{aligned}$$

Comparing to the case with $\left\{\widehat{\psi}_{t^{**}} < \underline{\psi}\right\}$, we see that the term in the expectation is identical, but for the coefficient of $\widehat{\psi}_{t^{**}}$, which is multiplied by $(1-\gamma)$. The distribution of the exponent is given in Lemma A6 for $c_1 = \gamma$. In this case, defining

$$y_{H,t^{**}} = -\gamma b_{t^{**}} + b_{t^{**}}^N - \gamma A_1(\tau^{**}) \rho_{t^{**}} + A_1(\tau^{**}) \rho_{t^{**}}^N + (1-\gamma) A_2(\tau^{**}) \widehat{\psi}_{t^{**}}$$

we have that

$$\mu_{H,y,t} = E[y_{H,t^{**}}] = -\gamma b_t + b_t^N + a_0(t) - \gamma A_1(\tau) \rho_t + A_1(\tau) \rho_t^N + (A_2(\tau) - \gamma A_2(\tau^{**})) \widehat{\psi}_t$$

The same steps then show

$$\begin{aligned} E_t\left[e^{y_{H,t^{**}}} | \widehat{\psi}_{t^{**}} > \underline{\psi}\right] &= \frac{1}{1 - N\left(\underline{\psi}; \widehat{\psi}_t, \sigma_{\widehat{\psi}}^2\right)} \int_{\underline{\psi}}^{\infty} E\left[e^{y_{H,t^{**}}} | \widehat{\psi}_{t^{**}}\right] f\left(\widehat{\psi}_{t^{**}}\right) d\widehat{\psi}_{t^{**}} \\ &= B_t^{-\gamma} B_t^N e^{a_0(t) - \gamma A_1(t;T)\rho_t + A_1(t;T)\rho_t^N + (A_2(\tau;T) - \gamma A_2(t^{**};T))\widehat{\psi}_t + \frac{1}{2}\sigma_{Hy}^2} R_{H,t}^N \end{aligned}$$

where $R_{H,t}^N$ is defined in Proposition 5.

So, we finally obtain

$$\begin{aligned} E_t\left[B_T^{-\gamma} B_T^N | \widehat{\psi}_{t^{**}} \ge \underline{\psi}\right] &= B_t^{-\gamma} B_t^N (1-\kappa)^{-\gamma} e^{C_0(\tau^{**}) + \frac{1}{2}(1-\gamma)^2 A_2(\tau^{**})^2 \widehat{\sigma}_{t^{**}}^2} \\ &\quad \times e^{a_0(t) - \gamma A_1(\tau)\rho_t + A_1(\tau)\rho_t^N + (A_2(\tau) - \gamma A_2(\tau^{**}))\widehat{\psi}_t + \frac{1}{2}\sigma_{Hy}^2} R_{H,t}^N \\ &= B_t^{-\gamma} B_t^N (1-\kappa)^{-\gamma} e^{C_0(\tau) - \gamma A_1(\tau)\rho_t + A_1(\tau)\rho_t^N + (A_2(\tau) - \gamma A_2(\tau^{**}))\widehat{\psi}_t + \frac{1}{2}(A_2(\tau) - \gamma A_2(\tau^{**}))^2 \widehat{\sigma}_t^2} R_{H,t}^N \end{aligned}$$

where the second equality is obtained from the first after some tedious algebra. Putting all terms together, we obtain the expression in Proposition 5. Q.E.D.

Proof of Corollary 2: The proof follows from an application of Ito's Lemma to the respective pricing functions. We obtain

$$\boldsymbol{\sigma}_M^N = A_1(\tau)\boldsymbol{\sigma}_N + \left(S_{M,t}^N + S_{\pi,t}\right) \widetilde{\boldsymbol{\sigma}}_\psi$$

where $\boldsymbol{\sigma}_N = (\sigma_{N,0}, \sigma_{N,1})$ and

$$S^N_{M,t} = \frac{\left(A_2(\tau) + \frac{1}{p^N_{L,t}}\frac{\partial p^N_{L,t}}{\partial \widehat{\psi}}\right) K^{no}_t + \left((A_2(\tau) - \gamma A_2(\tau^{**})) + \frac{1}{p^N_{H,t}}\frac{\partial p^N_{H,t}}{\partial \widehat{\psi}}\right) K^{yes}_t}{(1-p_t) K^{no}_t + p_t K^{yes}_t} \tag{B27}$$

with

$$p^N_{L,t} \equiv p^N_L\left(\widehat{\psi}_t, t\right) = \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + \sigma^L_{y\widehat{\psi}}, \sigma^2_{\widehat{\psi},t}\right) \tag{B28}$$

$$p^N_{H,t} \equiv p^N_H\left(\widehat{\psi}_t, t\right) = 1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + \sigma^H_{y\widehat{\psi}}, \sigma^2_{\widehat{\psi},t}\right) \tag{B29}$$

For the old economy

$$\boldsymbol{\sigma}_M = A_1(\tau)\boldsymbol{\sigma} + (S_{M,t} + S_{\pi,t})\,\widetilde{\boldsymbol{\sigma}}_\psi$$

where

$$S_{M,t} = \frac{-\frac{\partial p}{\partial \widehat{\psi}} G^{no}_t + \left((1-\gamma) A_2(\tau^{**}) + \frac{1}{\overline{p}}\frac{\partial \overline{p}}{\partial \widehat{\psi}}\right) G^{yes}}{(1-p_t) G^{no}_t + p_t G^{yes}_t} \tag{B30}$$

and

$$\overline{p}_t \equiv \overline{p}\left(\widehat{\psi}_t, t\right) = 1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + (1-\gamma) A_2(\tau^{**})\sigma^2_{\widehat{\psi},t}, \sigma^2_{\widehat{\psi},t}\right)$$

Q.E.D.

Proof of Proposition 6: Consider the old economy first. Rewrite the M/B of the old economy as

$$MB_t = \frac{G^{no}_t + p_t H_t}{\widetilde{G}^{no}_t + p_t \widetilde{H}_t}$$

where $H_t = G^{yes}_t - G^{no}_t$, and $\widetilde{H}_t = \widetilde{G}^{yes}_t - \widetilde{G}^{no}_t$. Given the closed form formulas for all the functions, we can compute the first derivative of MB_t with respect to the probability of adoption of the new technology p_t :

$$\frac{\partial MB_t}{\partial p_t} = \frac{H_t \widetilde{G}^{no}_t - G^{no}_t \widetilde{H}_t}{\left(\widetilde{G}^{no}_t + p_t \widetilde{H}_t\right)^2}$$

That is, the M/B increases in p_t if and only if $H_t \widetilde{G}^{no}_t > G^{no}_t \widetilde{H}_t$. Substituting the closed form expressions, we obtain the condition $h_{old} > 0$ where

$$h_{old} = -\widetilde{\kappa} + A_2(\tau^{**})\,\widehat{\psi}_t + \frac{1}{2}(1-2\gamma) A_2(\tau^{**})^2 \widehat{\sigma}^2_t \tag{B31}$$

$$- \log\left(\frac{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2(\tau^{**})\sigma^2_{\widehat{\psi},t}, \sigma^2_{\widehat{\psi},t}\right)}{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + (1-\gamma) A_2(\tau^{**})\sigma^2_{\widehat{\psi},t}, \sigma^2_{\widehat{\psi},t}\right)}\right) \tag{B32}$$

Consider now the new economy

$$MB^N_t = \frac{K_t R^N_{L,t} + p_t J_t}{\widetilde{G}^{no}_t + p_t \widetilde{H}_t}$$

where $\overline{J}_t = (1-\kappa)^{-\gamma} K_t^N R_H^N - K_t R_L^N$. The first derivative with respect to p_t is

$$\frac{\partial MB_t^N}{\partial p_t} = \frac{\overline{J}_t \widetilde{G}_t^{no} - K_t R_{L,t}^N \widetilde{H}_t}{\left(\widetilde{G}_t^{no} + p_t \widetilde{H}_t\right)^2}$$

Once again, the M/B of the new economy increases in p_t if and only if $\overline{J}_t \widetilde{G}_t^{no} - K_t \widetilde{H}_t R_{L,t}^N > 0$. Substituting, we obtain the condition $h_{new} > 0$

$$h_{new} = -\gamma A_2 (\tau^{**}) A_2 (\tau) \widehat{\sigma}_t^2 - \log \left(\frac{\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + \sigma^L_{y\widehat{\psi}}, \sigma^2_{\widehat{\psi},t}\right)}{\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \sigma^2_{\widehat{\psi},t}\right)} \right) \tag{B33}$$

$$- \log \left(\frac{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2 (\tau^{**}) \sigma^2_{\widehat{\psi},t}, \sigma^2_{\widehat{\psi},t}\right)}{1-\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2 (\tau^{**}) \sigma^2_{\widehat{\psi},t} + \sigma^L_{y\psi}, \sigma^2_{\widehat{\psi},t}\right)} \right) \tag{B34}$$

Proof of Proposition 7: Consider $\frac{M^N}{B^N} = \frac{\Phi^N}{\widetilde{\pi}}$, where Φ^N and $\widetilde{\pi}$ are defined appropriately. Then,

$$\frac{\partial \left(\frac{M^N}{B^N}\right)}{\partial \widehat{\psi}_t} = \frac{\widetilde{\pi} \partial \Phi^N / \partial \widehat{\psi}_t - \Phi^N \partial \widetilde{\pi} / \partial \widehat{\psi}_t}{\widetilde{\pi}^2} > 0$$

if and only if $S^N_{M,t} + S_{\pi,t} > 0$ where $S^N_{M,t}$, and $S_{\pi,t}$ are defined above. The probability of adoption as of time t^* is given by

$$p_{t^*} = \int_{f\left(\kappa,\gamma,\widehat{\sigma}^2_{t^*};\tau^*\right)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

where

$$f\left(\kappa,\gamma,\widehat{\sigma}^2_{t^*};\tau^*\right) = -\log(1-\kappa)/A_2(\tau^{**}) \left(\frac{\left(\widehat{\sigma}^2_{t^*}\right)^{-1} + \left(\frac{\phi}{\sigma_{N,1}}\right)^2 (t^{**}-t^*)}{\widehat{\sigma}^2_{t^*} \left(\frac{\phi}{\sigma_{N,1}}\right)^2 (t^{**}-t^*)} \right)^{\frac{1}{2}}$$

$$+ \frac{\frac{1}{2}(\gamma-1) A_2(\tau^{**})}{\left(\frac{\phi}{\sigma_{N,1}}\right) (t^{**}-t^*)^{\frac{1}{2}} \left(1 + \widehat{\sigma}^2_{t^*} \left(\frac{\phi}{\sigma_{N,1}}\right)^2 (t^{**}-t^*)\right)^{\frac{1}{2}}}$$

Thus, p_t is small whenever $f\left(\kappa,\gamma,\widehat{\sigma}^2_{t^*};\tau^*\right)$ is large. We can see that $f\left(\kappa,\gamma,\widehat{\sigma}^2_{t^*};\tau^*\right)$ is large when κ is high, γ is high and, finally, when $\widehat{\sigma}^2_{t^*}$ is small (if $\kappa > 0$). (In addition, we can see that f is large when $(t^{**}-t^*)$ is small and T is large, the latter due to the increase in $A_2(\tau^{**}) = (T-t^{**}) - \left(1-e^{-\phi(T-t^{**})}\right)/\phi$). In all of these cases, the formulas for the various quantities in $S^N_{M,t} + S_{\pi,t}$ imply that the latter becomes positive. Q.E.D.

Proof of Lemma A6: Let

$$y_{t^{**}} = -\gamma b_{t^{**}} + b^N_{t^{**}} - \gamma A_1(t^{**};T)\rho_{t^{**}} + A_1(t^{**};T)\rho^N_{t^{**}} + (1-c_1) A_2(t^{**};T) \widehat{\psi}_{t^{**}}$$

The fact that $y_{t^{**}}$ and $\widehat{\psi}_{t^{**}}$ are jointly normally distributed stems from the linearity of all of the processes. To compute the means, variances and covariances, we can compute the joint moment generating function. That is, let $\alpha_1, \alpha_2 > 0$, and define

$$N\left(b_t, b_t^N, \rho_t, \rho_t^N, \widehat{\psi}_t, \widehat{\sigma}_t^2, t\right) = E_t\left[e^{\alpha_1 y_{t^{**}} + \alpha_2 \widehat{\psi}_{t^{**}}}\right]$$

where the processes of stochastic variables are given by (B7) - (B12) with $c = 0$. Let $\mathbf{x}_t = \left(b_t, b_t^N, \rho_t, \rho_t^N, \widehat{\psi}_t, \widehat{\sigma}_t^2\right)$, the Feynman-Kac theorem shows that N must satisfy the PDE

$$0 = \frac{\partial N}{\partial t} + \sum_i \frac{\partial N}{\partial x_i} E_t[dx] + \frac{1}{2}\sum_i \sum_j \frac{\partial^2 N}{\partial x_i \partial x_j} E_t[dx_i dx_j]$$

with the boundary condition $N\left(b_{t^{**}}, b_{t^{**}}^N, \rho_{t^{**}}, \rho_{t^{**}}^N, \widehat{\psi}_{t^{**}}, \widehat{\sigma}_{t^{**}}^2, t^{**}\right) = e^{\alpha_1 y_{t^{**}} + \alpha_2 \widehat{\psi}_{t^{**}}}$. It can be verified that the solution to the PDE is given by

$$N_t = e^{\alpha_1\left\{-\gamma b_t + b_t^N - \gamma C_1(t;T)\rho_t + C_1(t;T)\rho_t^N\right\} + \alpha_1 C_0(t;T) + \{(1-c_1)\alpha_1 C_2(t;T) + \alpha_2\}\widehat{\psi}_t + \alpha_1 C_3(t;T)\widehat{\sigma}_t^2}$$

where

$$\begin{aligned}
C_1(t;T) &= \frac{1 - e^{-\phi(T-t)}}{\phi} = A_1(t;T) \\
C_2(t;T) &= A_2(t^{**};T) + \frac{1}{(1-c_1)} a(t) \\
\alpha_1 C_3(t;T) &= \widetilde{C}_3(t;T) \\
&= \frac{1}{2}\left((1-c_1)\alpha_1 C_2 + \alpha_2\right)^2 - \frac{1}{2}\left((1-c_1)\alpha_1 A_2(t^{**};T) + \alpha_2\right)^2
\end{aligned}$$

and

$$\begin{aligned}
\alpha_1 C_0(t;T) &= \widetilde{C}_0(t;T) \\
&= \alpha_1(1-\gamma)\overline{\rho} a(t) + \frac{1}{2}\left((1-c_1)\alpha_1 A_2(t^{**};T) + \alpha_2\right)^2\left[\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right] \\
&\quad + \alpha_1^2 \frac{1}{2}(\sigma^*)^2 a_2(t)
\end{aligned}$$

Above, $a(t)$, $a_2(t)$ and σ^* are given by (B23) - (B25). Rewrite $N_t = e^{g(\alpha_1, \alpha_2)}$ where

$$g(\alpha_1, \alpha_2) = \alpha_1\left\{-\gamma b_t + b_t^N - \gamma C_1(t;T)\rho_t + C_1(t;T)\rho_t^N\right\} + \widetilde{C}_0(t;T) + \{(1-c_1)\alpha_1 C_2(t;T) + \alpha_2\}\widehat{\psi}_t + \widetilde{C}_3(t;T)\widehat{\sigma}_t^2$$

Thus

$$\frac{\partial N}{\partial \alpha_1} = e^g\left\{\left(-\gamma b_t + b_t^N - \gamma C_1 \rho_t + C_1 \rho_t^N\right) + \frac{\partial \widetilde{C}_0}{\partial \alpha_1} + (1-c_1) C_2 \widehat{\psi}_t + \frac{\partial \widetilde{C}_3}{\partial \alpha_1}\widehat{\sigma}_t^2\right\}$$

We can use

$$\begin{aligned}
\frac{\partial \widetilde{C}_0}{\partial \alpha_1} &= (1-\gamma)\overline{\rho} a(t) + \left((1-c_1)\alpha_1 A_2(t^{**};T) + \alpha_2\right)(1-c_1) A_2(t^{**};T)\left[\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right] \\
&\quad + \alpha_1 (\sigma^*)^2 a_2(t)
\end{aligned}$$

and

$$\frac{\partial \widetilde{C}_3}{\partial \alpha_1} = ((1-c_1)\,\alpha_1 C_2 + \alpha_2)\,(1-c_1)\,C_2 - ((1-c_1)\,\alpha_1 A_2\,(t^{**};T) + \alpha_2)\,(1-c_1)\,A_2\,(t^{**};T)$$

Thus

$$\lim_{\alpha_1,\alpha_2\to 0}\frac{\partial N}{\partial \alpha_1} = \mu_y = \left\{-\gamma b_t + b_t^N - \gamma C_1\rho_t + C_1\rho_t^N + (1-\gamma)\,\overline{\rho} a\,(t) + (1-c_1)\,C_2\widehat{\psi}_t\right\}$$

Similarly

$$\frac{\partial N}{\partial \alpha_2} = e^g\left\{\frac{\partial \widetilde{C}_0}{\partial \alpha_2} + \widehat{\psi}_t + \frac{\partial \widetilde{C}_3}{\partial \alpha_2}\widehat{\sigma}_t^2\right\}$$

Since

$$\begin{aligned}\frac{\partial \widetilde{C}_0}{\partial \alpha_2} &= ((1-c_1)\,\alpha_1 A_2\,(t^{**};T) + \alpha_2)\left[\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right] \\ \frac{\partial \widetilde{C}_3}{\partial \alpha_2} &= ((1-c_1)\,\alpha_1 C_2 + \alpha_2) - ((1-c_1)\,\alpha_1 A_2\,(t^{**};T) + \alpha_2)\end{aligned}$$

we find

$$\lim_{\alpha_1,\alpha_2\to 0}\frac{\partial N}{\partial \alpha_2} = \mu_\psi = \widehat{\psi}_t$$

Turning to the second moments

$$\begin{aligned}\frac{\partial^2 N}{\partial \alpha_1^2} &= e^g\left\{\left(-\gamma b_t + b_t^N - \gamma C_1\rho_t + C_1\rho_t^N\right) + \frac{\partial \widetilde{C}_0}{\partial \alpha_1} + (1-c_1)\,C_2\widehat{\psi}_t + \frac{\partial \widetilde{C}_3}{\partial \alpha_1}\widehat{\sigma}_t^2\right\}^2 \\ &\quad + e^g\left\{\frac{\partial^2 \widetilde{C}_0}{\partial \alpha_1^2} + \frac{\partial^2 \widetilde{C}_3}{\partial \alpha_1^2}\widehat{\sigma}_t^2\right\}\end{aligned}$$

Since

$$\begin{aligned}\frac{\partial^2 \widetilde{C}_0}{\partial \alpha_1^2} &= (1-c_1)^2\,A_2\,(t^{**};T)^2\left[\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right] + (\sigma^*)^2\,a_2\,(t) \\ \frac{\partial^2 \widetilde{C}_3}{\partial \alpha_1^2} &= ((1-c_1)\,C_2)^2 - ((1-c_1)\,A_2\,(t^{**};T))^2\end{aligned}$$

we obtain

$$\begin{aligned}\lim_{\alpha_1,\alpha_2\to 0}\frac{\partial^2 N}{\partial \alpha_1^2} &= \left\{-\gamma b_t + b_t^N - \gamma C_1\rho_t + C_1\rho_t^N + (1-\gamma)\,\overline{\rho} a\,(t) + (1-c_1)\,C_2\widehat{\psi}_t\right\}^2 \\ &\quad + (1-c_1)^2\,A_2\,(t^{**};T)^2\left[\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right] + (\sigma^*)^2\,a_2\,(t) \\ &\quad + \left(((1-c_1)\,C_2)^2 - ((1-c_1)\,A_2\,(t^{**};T))^2\right)\widehat{\sigma}_t^2\end{aligned}$$

Thus

$$\begin{aligned}\sigma_y^2 &= \lim_{\alpha_1,\alpha_2\to 0}\frac{\partial^2 N}{\partial \alpha_1^2} - \left(\lim_{\alpha_1,\alpha_2\to 0}\frac{\partial N}{\partial \alpha_1}\right)^2 \\ &= (1-c_1)^2\,A_2\,(t^{**};T)^2\left(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right) + a\,(\tau)^2\,\widehat{\sigma}_t^2 + 2A_2\,(t^{**};T)\,(1-c_1)\,a\,(\tau)\,\widehat{\sigma}_t^2 + (\sigma^*)^2\,a_2\,(t)\end{aligned}$$

Similarly,

$$\frac{\partial^2 N}{\partial \alpha_2^2} = e^g \left\{ \frac{\partial \widetilde{C}_0}{\partial \alpha_2} + \widehat{\psi}_t + \frac{\partial \widetilde{C}_3}{\partial \alpha_2} \widehat{\sigma}_t^2 \right\}^2 + e^g \left\{ \frac{\partial^2 \widetilde{C}_0}{\partial \alpha_2^2} + \frac{\partial^2 \widetilde{C}_3}{\partial \alpha_2^2} \widehat{\sigma}_t^2 \right\}$$

Since

$$\begin{aligned} \frac{\partial^2 \widetilde{C}_0}{\partial \alpha_2^2} &= \left[\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2 \right] \\ \frac{\partial \widetilde{C}_3}{\partial \alpha_2} &= 0 \end{aligned}$$

we have

$$\lim_{\alpha_1, \alpha_2 \to 0} \frac{\partial^2 N}{\partial \alpha_2^2} = \widehat{\psi}_t^2 + \left[\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2 \right]$$

and thus

$$\sigma_\psi^2 = \lim_{\alpha_1, \alpha_2 \to 0} \frac{\partial^2 N}{\partial \alpha_2^2} - \left(\lim_{\alpha_1, \alpha_2 \to 0} \frac{\partial N}{\partial \alpha_2} \right)^2 = \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2$$

Finally

$$\begin{aligned} \frac{\partial^2 N}{\partial \alpha_2 \partial \alpha_1} &= e^g \left\{ \left(-\gamma b_t + b_t^N - \gamma C_1 \rho_t + C_1 \rho_t^N \right) + \frac{\partial \widetilde{C}_0}{\partial \alpha_1} + (1 - c_1) C_2 \widehat{\psi}_t + \frac{\partial \widetilde{C}_3}{\partial \alpha_1} \widehat{\sigma}_t^2 \right\} \left\{ \frac{\partial \widetilde{C}_0}{\partial \alpha_2} + \widehat{\psi}_t + \frac{\partial \widetilde{C}_3}{\partial \alpha_2} \widehat{\sigma}_t^2 \right\} \\ &\quad + e^g \left\{ \frac{\partial^2 \widetilde{C}_0}{\partial \alpha_2 \partial \alpha_1} + \frac{\partial^2 \widetilde{C}_3}{\partial \alpha_2 \partial \alpha_1} \widehat{\sigma}_t^2 \right\} \end{aligned}$$

Since

$$\begin{aligned} \frac{\partial^2 \widetilde{C}_0}{\partial \alpha_2 \partial \alpha_1} &= (1 - c_1) A_2 (t^{**}; T) \left[\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2 \right] \\ \frac{\partial^2 \widetilde{C}_3}{\partial \alpha_2 \partial \alpha_1} &= (1 - c_1) (C_2 - A_2 (t^{**}; T)) \end{aligned}$$

we have

$$\begin{aligned} \lim_{\alpha_1, \alpha_2 \to 0} \frac{\partial^2 N}{\partial \alpha_2 \partial \alpha_1} &= \left\{ -\gamma b_t + b_t^N - \gamma C_1 \rho_t + C_1 \rho_t^N + (1 - \gamma) \overline{\rho} a(t) + (1 - c_1) C_2 \widehat{\psi}_t \right\} \left\{ \widehat{\psi}_t \right\} \\ &\quad + \left\{ (1 - c_1) A_2 (t^{**}; T) \left[\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2 \right] + (1 - c_1) (C_2 - A_2 (t^{**}; T)) \widehat{\sigma}_t^2 \right\} \end{aligned}$$

implying

$$\begin{aligned} \sigma_{y,\psi} &= \left(\lim_{\alpha_1, \alpha_2 \to 0} \frac{\partial^2 N}{\partial \alpha_2 \partial \alpha_1} \right) - \left(\lim_{\alpha_1, \alpha_2 \to 0} \frac{\partial N}{\partial \alpha_1} \right) \left(\lim_{\alpha_1, \alpha_2 \to 0} \frac{\partial N}{\partial \alpha_2} \right) \\ &= (1 - c_1) A_2 (t^{**}; T) \left(\widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2 \right) + a(t) \widehat{\sigma}_t^2 \end{aligned}$$

Some additional algebra yields the formulas in Lemma A6. Q.E.D.

REFERENCES

Agarwal, Deepak, Sreedhar T. Bharath, and Siva Viswanathan, 2004, Technological change and stock return volatility: Evidence from eCommerce adoptions, Working paper, University of Michigan.

Aghion, Philippe, and Peter Howitt, 1992, A model of growth through creative destruction, *Econometrica* 60, 323–351.

Arrow, Kenneth J., 1962, The economic implications of learning by doing, *Review of Economic Studies* 29, 155-173.

Atkeson, Andrew, and Patrick J. Kehoe, 2003, The Transition to a New Economy After the Second Industrial Revolution, Working paper, Federal Reserve Bank of Minneapolis.

Benmelech, Efraim, 2005, Asset salability and debt maturity: Evidence from 19th century American railroads, Working paper, Harvard University.

Brown, Stephen J., William N. Goetzmann, and Stephen A. Ross, 1995, Survival, *Journal of Finance* 50, 853–873.

Caselli, Francesco, 1999, Technological revolutions, *American Economic Review* 89, 78–102.

Chari, V.V., and Hugo Hopenhayn, 1991, Vintage human capital, growth, and the diffusion of new technology, *Journal of Political Economy* 99, 1142–1165.

David, Paul A., 1991, Computer and dynamo: The modern productivity paradox in a not-too-distant mirror, In *Technology and Productivity: The Challenge for Economic Policy* (OECD, Paris), 315–347.

Fogel, Robert W., 1964, *Railroads and American Economic Growth: Essays in Econometric History*, Johns Hopkins Press, Baltimore, MD.

Goetzmann, William N., Roger G. Ibbotson, and Liang Peng, 2001, A new historical database for the NYSE 1815 to 1925: Performance and predictability, *Journal of Financial Markets* 4, 1–32.

Greenwood, Jeremy, and Boyan Jovanovic, 1999, The information-technology revolution and the stock market, *American Economic Review Papers and Proceedings* 89, 116–122.

Hobijn, Bart, and Boyan Jovanovic, 2001, The Information-technology revolution and the stock market: Evidence, *American Economic Review* 91, 1203–1220.

Johnson, Timothy C., 2005, Optimal learning and new technology bubbles, *Journal of Monetary Economics*, forthcoming.

Jovanovic, Boyan, 1982, Selection and the evolution of industry, *Econometrica* 50, 649–670.

Jovanovic, Boyan, and Glenn M. MacDonald, 1994, The life cycle of a competitive industry, *Journal of Political Economy* 102, 322–347.

Jovanovic, Boyan, and Yaw Nyarko, 1996, Learning by doing and the choice of technology, *Econometrica* 64, 1299-1310.

Jovanovic, Boyan, and Peter L. Rousseau, 2003, Two technological revolutions, *Journal of European Economic Association* 1, 419–428.

Jovanovic, Boyan, and Peter L. Rousseau, 2005, General purpose technologies, NBER Working paper #11093.

Klein, Maury, 1994, *Unfinished Business: The Railroad in American Life*, University Press of New England.

Laitner, John, and Dmitriy Stolyarov, 2003, Technological change and the stock market, *American Economic Review* 93, 1240–1267.

Laitner, John, and Dmitriy Stolyarov, 2004a, Aggregate returns to scale and embodied technical change: Theory and measurement using stock market data, *Journal of Monetary Economics* 51, 191–233.

Laitner, John, and Dmitriy Stolyarov, 2004b, The role of owned ideas in stock market runups, Working paper, University of Michigan.

Malkiel, Burton G., 1999, *A Random Walk Down Wall Street*, W. W. Norton & Co., New York, NY.

Manuelli, Rodolfo E., 2003, Technological change, the labor market and the stock market, Working paper, University of Wisconsin-Madison.

Mazzucato, Mariana, 2002, The PC industry: New economy or early life-cycle?, *Review of Economic Dynamics* 5, 318–345.

Mishkin, Frederic S., 1991, Asymmetric information and financial crises: A historical perspective, In *Financial Markets and Financial Crises*, edited by R. Glenn Hubbard, University of Chicago Press, Chicago, 69–108.

Mokyr, Joel, 1990, *The Level of Riches: Technological Creativity and Economic Progress*, Oxford University Press, New York.

Pástor, Ľuboš, and Pietro Veronesi, 2003, "Stock valuation and learning about profitability," *Journal of Finance* 58, 1749–1789.

Pástor, Ľuboš, and Pietro Veronesi, 2006, "Was there a Nasdaq bubble in the late 1990s?" *Journal of Financial Economics*, forthcoming.

Perez, Carlota, 2002, *Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages*, Edward Elgar, Cheltenham, UK.

Romer, Paul, 1990, Endogenous technological change, *Journal of Political Economy* 98, S71–S102.

Shiller, Robert, 2000, *Irrational Exuberance*, Princeton University Press, Princeton, NJ.

Stover, John F., 1961, *American Railroads*, University of Chicago Press.

The Economist, 2000, September 21, Bubble.com.