Technological Revolutions and Stock Prices

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

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

February 7, 2007

Abstract

During technological revolutions, stock prices of innovative firms tend to exhibit bubble-like patterns. We develop a general equilibrium model that can rationalize these patterns. The average productivity of a new technology is uncertain, and investors learn about it before deciding whether to adopt the technology on a large scale. For technologies that are ultimately adopted, the nature of the uncertainty changes from idiosyncratic to systematic as the adoption becomes more likely, so stock prices fall after an initial run-up. The resulting "bubbles" are observable ex post but unpredictable ex ante, and they are most pronounced for technologies characterized by high uncertainty and fast adoption. We examine stock prices in 1830–1861 and 1992–2005 when the railroad and Internet technologies spread in the United States, and we find support for the model's predictions.

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 the audience participants at the 2007 American Finance Association meeting, the 2006 European Finance Association meeting, the 2006 UBC Summer Conference, the Fall 2005 NBER Asset Pricing meeting, and especially the conference discussants, Markus Brunnermeier, Leonid Kogan, Lars Lochstoer, and Lu Zhang, for helpful comments. We also thank Malcolm Baker, Robert Barro, Efraim Benmelech, Gene Fama, Bob Fogel, Boyan Jovanovic, John Heaton, Ali Lazrak, Robert Novy-Marx, Rob Stambaugh, Dmitriy Stolyarov, and the audiences at CERGE-EI, Dartmouth College, Ente Einaudi, Harvard University, Indiana University, London Business School, London School of Economics, New York University, Stockholm Institute for Financial Research, Stockholm School of Economics, University of Chicago, University of Pennsylvania, and University of Vienna. Shastri Sandy has provided valuable research assistance.

"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). 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 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 the 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.

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 a

[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] "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 Economist, September 21, 2000, Bubble.com)

large-scale 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 new technology is mostly idiosyncratic, the new economy stocks command high valuation ratios. As the adoption probability increases, the resulting increase in systematic risk pushes up the discount rates and thus depresses stock prices in both the new and old economies. The new economy stock prices fall deeper because their discount rates rise higher due to an increase in the new economy's market beta. 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 discount rates but also by expected 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 prevails 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 an irrational bubble but it obtains under rational expectations through a general equilibrium effect.

The bubble-like pattern in stock prices arises due to an ex post selection 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 rare 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.

In addition to the level of stock prices, the high stock return 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. After an initial decline, the new economy's volatility rises sharply when the stochastic discount factor becomes more volatile as a result of a higher probability of a large-scale adoption. The same effect also pushes up the new economy's market beta and the old economy's volatility, two different measures of systematic risk in the economy.

Our model makes many empirical predictions for technological revolutions: The "bubble" in stock prices should be much stronger in the new economy than in the old economy; stock prices in both economies should reach the bottom at the end of the revolution; the new economy's market beta should increase sharply before the end of the revolution; the new economy's volatility should also rise sharply and it should exceed the old economy's volatility; the old economy's volatility should rise but less than the new economy's volatility; the new economy's beta and both volatilities should all peak at the end of the revolution; and the old economy's productivity should begin rising at the end of the revolution.

All of these predictions are supported by the empirical evidence from the recent Internet revolution. According to the model, this revolution ended (i.e., the probability of a large-scale adoption of the Internet technology reached one) in 2002. The "bubble" pattern was much stronger in the NASDAQ index (our proxy for the new economy) than in the NYSE/AMEX index (the old economy); both stock price indexes reached the bottom in 2002; NASDAQ's beta doubled between 1997 and 2002; NYSE/AMEX's return volatility also doubled and NASDAQ's volatility tripled over the same period; NASDAQ's volatility always exceeded NYSE/AMEX's volatility; NASDAQ's beta and both volatilities peaked in 2002; and the productivity growth of the U.S. economy accelerated sharply after 2002.

We also examine stock prices during the first major technological revolution in the U.S. since the opening of the U.S. stock market – the introduction of steam-powered railroads. 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. The railroad stock volatility and price-dividend ratios consistently exceeded their non-railroad counterparts. The volatility of all stocks rose in 1857. The railroad stock beta increased sharply in the 1850s, before falling right after 1857. In the context of our model, all of 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). Different but related models of learning are presented in Jovanovic (1982), Jovanovic and Nyarko (1996), and Atkeson and Kehoe (2006). We empirically examine the Internet and railroad revolutions, 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.

The paper is organized as follows. Section 2 presents the model. Section 3 solves for stock prices and analyzes their dynamics. Section 4 investigates the model's empirical predictions for stock prices during technological revolutions. Section 5 empirically examines the behavior of stock prices in 1830–1861 and 1992–2005 when the railroad technology and the Internet technology, respectively, 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

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

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

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 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 $\kappa \geq 0$.

The agent has three choices at time t^* when the new technology becomes available:

(i) Adopt the new technology

(ii) Begin learning about the new technology (i.e., about ψ)

(iii) Discard the new technology

We show that the agent optimally chooses option (ii). Learning is described in Section 2.1.

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.[3]

[3] 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.

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

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

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

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

where the left-hand side is expected utility conditional on adopting the technology at time t^*, and the right-hand side is expected utility conditional on not adopting the technology at time t^* or any time afterwards.[4] The expected utility from no adoption at time t^* exceeds the right-hand side (and hence also the left-hand side) because it includes the value of the option to adopt after time t^*. Proposition 1 holds for any κ, including $\kappa = 0$, as it is driven 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 ρ_t 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:

[4] On the right hand side, V is evaluated at $\widehat{\psi}_{t^*} = \widehat{\sigma}^2_{t^*} = 0$. 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 $\psi = 0$.

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

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

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

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

Moreover, the productivity processes can be rewritten as

$$d\rho_t = \phi\left(\overline{\rho} - \rho_t\right) dt + \sigma d\widetilde{Z}_{0,t} \tag{11}$$

$$d\rho^N_t = \phi\left(\overline{\rho} + \widehat{\psi}_t - \rho^N_t\right) dt + \sigma_{N,0} d\widetilde{Z}_{0,t} + \sigma_{N,1} d\widetilde{Z}_{1,t}, \tag{12}$$

where the orthogonalized Brownian motions $\left(\widetilde{Z}_{0,t}, \widetilde{Z}_{1,t}\right)$, which capture the agent's expectation errors, are given in the Appendix.

Note that the posterior variance $\widehat{\sigma}^2_t$ declines deterministically over time due to learning. After the adoption at time t^{**}, the agent continues to learn about ψ by observing ρ^N_t and ρ_t, but the old economy's profitability follows equation (4) rather than equation (3).

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 stopping time t^{**} numerically in Section 4.2.

Until Section 4.2., we focus on a simpler problem in which the meaning of t^{**} is slightly different: Instead of denoting the optimal (endogenously chosen) adoption time, t^{**} denotes an exogenously given time at which the agent irreversibly decides whether or not to adopt the new technology. This simpler problem admits a closed-form solution 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 dynamics obtained when t^{**} is endogenously chosen are very similar to those obtained here with an exogenous t^{**}.

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

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

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, uncertainty $\widehat{\sigma}_{t^{**}}$, and risk aversion γ, which is intuitive. Using our closed-form expression for the value function in equation (6), equation (13) follows from the optimality condition

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

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

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

where the maximization involves choosing whether or not to adopt the new technology at time t^{**}, following Proposition 2.[5] The Appendix provides an expression for $\mathcal{V}$ (Lemma A3), along with a proof that expected utility is higher when experimentation takes place:

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

[5]Note the difference between the value functions $\mathcal{V}$ in equation (15) and V in equation (6). Whereas $\mathcal{V}$ includes the value of the option to adopt the new technology at the future time t^{**}, V does not include such option value because it applies to settings in which the adoption decision has already been made.

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. Our history is full of examples of 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

[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.

by M_t and M_t^N, respectively, are given by the standard pricing formulas

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

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, $t^* \le t < t^{**}$, 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

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

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

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 $\widehat{\psi}_{t^{**}} \ge \underline{\psi}$, 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.

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 $\widehat{\sigma}_t^2$.

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

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

[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.

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

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

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

$$\mu^N_{M,t} = -\sigma^{N,0}_{M,t}\sigma^0_{\pi,t} - \sigma^{N,1}_{M,t}\sigma^1_{\pi,t}, \tag{30}$$

and the components of the return volatilities are

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

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

with $S_{M,t}$ and $S^N_{M,t}$ 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.

In an earlier version of this paper, we plot h_{new} and h_{old} as functions of $\widehat{\psi}_t$ for the baseline parameter values (we drop the figure here to save space). We find that the condition $h_{new} > 0$ is satisfied when $\widehat{\psi}_t$ is close to its initial value of zero, but the condition becomes violated as $\widehat{\psi}_t$ increases towards the threshold $\underline{\psi}$. That is, $h_{new} > 0$ holds shortly after time t^*, but it becomes violated as 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 during a technological revolution. The condition $h_{old} > 0$ is never satisfied for the baseline parameter values, so the old economy's M/B is always decreasing in p_t. When $\widehat{\psi}_t$ increases, h_{old} increases because adopting a new technology is more valuable when the technology is more productive. 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.

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 ($E_t[B_T]$ and $E_t\left[B^N_T\right]$) in both economies. 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 decline in M/B 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 $\widehat{\psi}_t$. 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.[10] The cash flow effect is large, though, because M^N/B^N in equation (28) is strongly increasing in $\widehat{\psi}_t$. 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 $\widehat{\psi}_t$ 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

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

[10] Analogously, if a stock option is deep out of the money, a small increase in the stock price does not change the option value by much since its delta is small and the option remains deep out of the money.

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

$$\frac{M^N_{t^{**}}}{B^N_{t^{**}}} = e^{\overline{C}_0(\tau^{**}) + A_1(\tau^{**})\rho^N_{t^{**}} + A_2(\tau^{**})\widehat{\psi}_{t^{**}} + \frac{1}{2}A_2(\tau^{**})^2\widehat{\sigma}^2_{t^{**}}}. \tag{34}$$

The M/B of the new economy is clearly lower when the technological revolution takes place. The reason is the uncertainty term $\widehat{\sigma}^2_t$, whose coefficient is negative in part (a) and positive in part (b). In part (a), $\widehat{\sigma}^2_t$ is systematic (it affects π_t), whereas in part (b), it is idiosyncratic (it does not affect π_t). Since $\widehat{\psi}_t$ (expected cash flow) is essentially the same in both scenarios, the difference between M/B in parts (a) and (b) is due to the discount rate effect. This knife-edge case shows that the discount rate effect is generally stronger than the cash flow effect close to the adoption time t^{**}, so that 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 the key variables during technological revolutions. We simulate 50,000 samples of shocks in our economy and compute the paths of quantities such 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 is 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 simple 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 2 plots the average paths of $\widehat{\psi}_t$, p_t, and $\sigma_\pi \equiv \text{Std}(d\pi_t/\pi_t)$. 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).

Figure 3 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 3 plots the average paths of M/B across all technological revolutions. The new economy's M/B rises and then falls, as predicted in Section 3.1. Since we are conditioning on the adoption of the new technology at time t^{**}, $\widehat{\psi}_t$ must go up between t^* and t^{**} (Figure 2). 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 a "bubble." For the old economy, the cash flow effect is weaker 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 decline in the old economy's M/B.

Different technological revolutions produce different paths of 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

[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.

[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, $p_{t^*} \approx 2\%$, as expected. In principle, any product innovation could potentially lead to a technological revolution, but very few do, both in reality and in our model.

of Figure 3. This average path shows that apparent bubbles are not merely possible in a rational world; they should in fact be expected during technological revolutions.

Panel B of Figure 3 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 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 3 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, so 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 plots the average return volatility across all no-revolution samples. In these samples, the discount rate effect is weak and volatility is roughly constant over time.

Panels A and B of Figure 4 plot the market beta of the new economy, β, defined as the slope from the regression of the new economy stock returns on the old economy stock returns. In Panel A, where we condition on $p_{t^{**}} = 1$ (revolution), β exhibits an asymmetric U-shape pattern, for the following reason. Positive shocks to $\widehat{\psi}_t$ always reduce the market value of the old economy stocks (because the discount rate effect always prevails in the old economy), but they increase the value of the new economy stocks initially while the cash flow effect prevails, leading to an initial decrease in β. Only after the discount rate effect overcomes the cash flow effect for the new economy, shocks to $\widehat{\psi}_t$ begin affecting the market values of both economies in the same direction, leading to an increase in β. Since the effect of $\widehat{\psi}_t$ on the old economy stocks increases with the adoption probability, the rise in β is more dramatic than the initial fall. After a mild decline in the first half of the revolution, β doubles in the second half, from 0.75 to 1.5. The average beta in the no-revolution samples, plotted in Panel B, is almost flat over time.

The previous two paragraphs describe two measures of systematic risk that increase during technological revolutions: the old economy's volatility and the new economy's beta. The increase in the old economy's volatility raises the discount rates for both economies, old and new, holding β constant. The increase in β gives an additional boost to the discount rate of the new economy (but not the old economy), so it is not surprising that stock prices fall by more in the new economy than in the old economy.

The remaining panels of Figure 4 plot the average realized returns (solid line) and expected returns (dashed line).[13] In technological revolutions, realized stock returns are first positive and then negative for both economies, due to an *ex post selection 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 be 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 due solely 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 old economy's market value to drop immediately at time t^* (Panel E of Figure 4). This drop is driven by two forces. The possibility of eventual adoption means that conversion costs might be paid at time t^{**}, and it also increases systematic risk and so drives up the discount rate.

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

4.1. Sensitivity Analysis

This section examines the sensitivity of the price dynamics to our parameter choices. Figure 5 is the counterpart of Panel A of Figure 3 (revolution), with various parameter changes.

In Panel A of Figure 5, risk aversion $\gamma = 3$, as opposed to $\gamma = 4$ in Figure 3. Lower risk aversion increases M/B in both economies, as expected, but the pattern of M/B is otherwise the same as that in Figure 3. A hump-shaped pattern in M^N/B^N obtains for any $\gamma > 1$.

In Panel B of Figure 5, the cost of switching to the new technology is $\kappa = 0$, as opposed to $\kappa = 0.1$ in Figure 3. 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 3.

In Panel C, prior uncertainty about ψ is $\widehat{\sigma}_{t^*} = 8\%$, compared to $\widehat{\sigma}_{t^*} = 4\%$ in Figure 3. 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 5, the time until the adoption decision is shortened to $t^{**} - t^* = 4$, compared to $t^{**} - t^* = 8$ in Figure 3. 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 3. M^N/B^N then rises higher and falls deeper than in Figure 3, 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 systematic. Since both effects are stronger, the rise and fall in M^N/B^N are more striking than in Figure 3. 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 (if any) 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 6 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 very different t^{**}'s would not be very meaningful. For better comparison with Figure 3 in which t^{**} is fixed at 9 years, the left panels of Figure 6 report averages across those simulations in which the optimal t^{**} is between years 8 and 10. Our main results are unaffected by endogenizing t^{**}. The new economy's M/B is lower than in Figure 3, mostly because the optimal t^{**} exceeds 9 years, on average, and slower adoption reduces M/B. More important, during revolutions, this M/B exhibits a rise-and-fall pattern similar to that in Figure 3, albeit slightly weaker (a clearer "bubble" pattern is obtained for $\gamma = 3$, as we show in the previous draft). The path of volatility in Panel C is also quite similar, except that the volatility spike observed in Figure 3 disappears, as argued earlier.

5. Empirical Evidence

In this section, we empirically examine the behavior of stock prices during two technological revolutions, one recent and one distant. For both revolutions, we consider the key quantities in our model, such as the new economy's market beta and the level and volatility of stock prices, and compare their empirical dynamics with their model-implied dynamics.

5.1. The Internet Revolution

The Internet's predecessor, Arpanet, was created in 1969 with funding from the U.S. Department of Defense. Arpanet ceased to exist in 1990, at about the same time that Tim Berners-Lee and his team at CERN released the World Wide Web. The first Web site, info.cern.ch, was put online in 1991. The first graphics-based web browser, Mosaic, was launched in 1993 by Marc Andreessen at the National Center for Supercomputing Applications. In 1994, Andreessen co-founded Netscape Communications, which went public in August 1995 in the first Internet IPO. The first big pioneer of e-commerce was the online bookseller Amazon.com, which was launched by Jeff Bezos in 1995 and went public in May 1997. The Internet gradually became mainstream. The number of web servers grew from about 23,000 in mid-1995 to about one million in mid-1997, six million in mid-1999, 30 million in mid-2001, 41 million in mid-2003, and 65 million in mid-2005 (see www.zakon.org/robert/internet/timeline/). A prominent example of the Internet's integration into traditional business models was the creation of the first "clicks-and-mortar" company through the merger of AOL and Time Warner.[14] Today, the Internet technology is an indelible part of the economic landscape.

To provide a benchmark for our empirical analysis, we plot the model-implied dynamics of some key variables in Figure 7. These are the expected dynamics during a revolution, in that we average the model-implied paths across many simulations in which the new technology is adopted at time t^{**}. We keep all parameters from the baseline case (Table 1) except that we shorten the duration of the revolution from eight to six years because the Internet revolution was relatively fast. Panel A of Figure 7 shows that the new economy's market beta decreases slightly (from 0.9 to 0.7) in the first half of the revolution, but then it increases sharply in the second half, reaching 1.65 at time t^{**} before falling to one. This increase in beta is even steeper than in the baseline case (Panel A of Figure 4). Panel B shows that the increase in stock return volatility is also steeper than in the baseline case (Panel C of Figure 3), with the old economy's volatility doubling to 38% and the new economy's volatility rising to over 65%. Panel C plots the market values of both economies. There is a clear "bubble" in the new economy, whose market value quintuples and then falls by half. The old economy's market value also rises and falls, but this pattern is much weaker than in the new economy. Panel D shows that the old economy's productivity begins rising immediately after the adoption of the new technology, when it begins mean-reverting toward a higher mean.

[14] AOL announced its plan to acquire Time Warner (for some $182bn in stock) in January 2000, and the FTC approved the deal in January 2001. "The merger, the largest deal in history, combines the nation's top internet service provider with the world's top media conglomerate. The deal also validates the Internet's role as a leader in the new world economy, while redefining what the next generation of digital-based leaders will look like." (CNN Money, Jan 10, 2000).

Figure 8 is an empirical counterpart of Figure 7 for the period 1992–2005. For simplicity, we assume that the technology-loaded NASDAQ index represents the new economy and the NYSE/AMEX index is the old economy. We obtain daily index returns from CRSP.

Panel A of Figure 8 plots the market beta of the NASDAQ index, along with its two-standard-error confidence bands. The beta is computed daily as the slope coefficient from the regression of the NASDAQ returns on the NYSE/AMEX returns over the most recent 500 trading days (i.e., about two years). After a slight decrease from about 1.2, NASDAQ's beta doubles from 1.0 to 2.0 between 1997 and mid-2002, and this increase is highly statistically significant. This empirical pattern is strikingly similar to the model-implied pattern in Panel A of Figure 7, in which the beta also decreases by about 0.2 before rising 2.3-fold by the end of the revolution. According to the model, the time when the beta peaks is the time of the large-scale adoption; hence the evidence on NASDAQ's beta is consistent with the probability of the Internet's large-scale adoption reaching one by mid-2002.

Panel B of Figure 8 plots the standard deviations of returns on the NASDAQ and NYSE/AMEX indices, computed daily over the most recent 500 trading days. NASDAQ's volatility falls from 17% in 1992 to 11% in 1995, before rising to 47% at the beginning of 2002. NYSE/AMEX's volatility falls from 13% in 1992 to 8% in 1995, before rising to 21% by the end of 2002. These patterns are similar to the model-implied patterns in Panel B of Figure 7 in several ways: (i) the new economy's volatility always exceeds the old economy's volatility; (ii) both volatilities generally rise over time, with a bit of a U-shape pattern; (iii) the new economy's volatility rises much faster; and (iv) both volatilities peak at about the same time. In the model, both volatilities peak at the time of the adoption; the volatility evidence is thus consistent with the Internet revolution ending sometime in 2002.

Panel C of Figure 8 plots the index levels for NASDAQ and NYSE/AMEX, namely, the value of $1 invested in these indices in January 1992, with dividend reinvestment. The NASDAQ index quadruples between 1996 and March 2000, but then it falls back to the 1996 level by October 2002, after which it rises again.[15] In contrast, NYSE/AMEX exhibits a much smaller rise and fall over the same period. This pattern is similar to the model-implied pattern in Panel C of Figure 7, in which the new economy's market value also exhibits a "bubble" but the old economy's rise and fall are much less pronounced. According to the model, the time when both indices stop falling is the time of the large-scale adoption; hence this evidence is consistent with the Internet's adoption by October 2002.

[15] Pástor and Veronesi (2006) argue that NASDAQ's valuation in March 2000 can be justified in a rational valuation model with plausible levels of uncertainty about the average profitability of NASDAQ firms.

Panel D of Figure 8 plots a three-year moving average of multifactor productivity growth in the private business sector of the U.S. economy. (This is the most commonly used multifactor productivity measure, according to the Bureau of Labor Statistics, which is the source of the data.) In year t, we plot the average annual productivity growth in years $t-2$, $t-1$, and t. Multifactor productivity growth averaged about 1% per year in the 1990s, but it increased sharply after year 2002: from 1% per year in 2002 to 1.5% in 2003 and 2.5% in 2004 and 2005. A similar pattern is observed for labor productivity.[16] The observed productivity pattern is similar to the model-implied pattern in Panel D of Figure 7, except that that figure plots the level of productivity as opposed to its growth rate.[17] In the model, the economy's productivity begins rising at the time of the adoption; hence the productivity evidence is consistent with a large-scale adoption of the Internet by 2002.

Overall, we find Figure 8 remarkably similar to Figure 7. The patterns of NASDAQ's beta and NYSE/AMEX's volatility show that both sectors experienced large increases in systematic risk in 1997–2002, supporting the key prediction of the model. To summarize, the empirical evidence seems consistent with the joint hypothesis that our model holds and that the Internet technology was adopted on a large scale by 2002.

5.2. American Railroads Before the Civil War

Our paper is motivated by the technological revolutions, listed in the introduction, that were accompanied by apparent bubbles in stock prices. In this section, we conduct an "out-of-sample" analysis of a revolution whose stock prices have not been analyzed before. 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

[16] In his Remarks Before Leadership South Carolina on August 31, 2006, Ben Bernanke argued that "One of the most important economic developments in the United States in the past decade or so has been a sustained increase in the growth rate of labor productivity... From the early 1970s until about 1995, productivity growth in the U.S. nonfarm business sector averaged about 1.5% per year... Between 1995 and 2000, however, the rate of productivity growth picked up significantly, to about 2.5% per year... Talk of the "new economy" faded with the sharp declines in the stock valuations of high-tech firms at the turn of the millennium. Yet, remarkably, productivity accelerated further in the early part of this decade. From the end of 2000 to the end of 2003, productivity rose at a 3.5% annual rate and it is estimated to have increased at an average annual rate of 2.25% since the end of 2003. These advances were achieved despite adverse developments that included the 2001 recession, the terrorist attacks of September 11, [etc.]."

[17] Our comparison seems reasonable because in the model, average productivity can grow only via technological revolutions, whereas in reality, there are also many non-revolutionary improvements in productivity. Therefore, in the data, it is the growth rate of productivity that sets a technological revolution apart.

the historical milestones of American RRs in Section 5.2.1., 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 5.2.2. In the context of our model, our evidence is consistent with a large-scale adoption of the RR technology around year 1857.

5.2.1. 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.[18] 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 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

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

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.[19] “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 9, 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

[19]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.

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.

5.2.2. 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 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 284 RR stocks (42.32% of the whole sample) by examining the stock names. 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.

We compute monthly RR (non-RR) index returns as price-weighted averages of monthly capital gains across all RR (non-RR) stocks.[20] 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 resulting return series can be viewed as approximations to the actual returns earned by investors.

Panel A of Figure 10 plots the beta of the RR index, with a two-standard-error confidence band. The beta in month t is the slope coefficient from the regression of the most recent 36 monthly RR returns (in months $t-35$ through t) on the non-RR returns. Not surprisingly, the beta estimates computed from only 36 observations have wide confidence bands. Nonetheless, it appears that the largest increase in beta took place in the 1850s: The beta estimate rose from about 0.2 in 1850 to about 1.8 in 1856, before falling to about 1.0 right after 1857. This empirical pattern is quite similar to the model-implied pattern in Panel A of Figure 4 if we assume that the RR technology was adopted on a large scale in 1857.

Panel B 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. The volatility

[20] 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.

difference is also due in part to the fact that the RR portfolio is less diversified than the non-RR portfolio, but it 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 B is that return volatility increases sharply in 1857, to 33.5% per year for RRs and to 23.1% for non-RRs. Comparison with Panel C of Figure 3 shows that both facts are consistent with a large-scale adoption of the RR technology in 1857.

Panel C plots the stock price index levels for the RR and non-RR industries, obtained by cumulating monthly returns in each industry. 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 shortly 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 earlier with the rise in interest rates, the stock market decline... and the widening of the interest rate spread." Mishkin's 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.

In Panel D of Figure 10, we do not plot productivity as we did in Figure 8 because, to our knowledge, productivity in this period has been computed only at a ten-year frequency from census data. We note, however, that the evidence points to a large increase in productivity after the late 1850s. For example, Cochrane (1979) argues that productivity advanced sharply just before the Civil War, and Craig and Weiss (1993) conclude that "the 1860s saw the greatest increase in output per farm worker of any decade in the 19th century." This evidence on productivity further strengthens the case for a large-scale adoption of the RR technology in the late 1850s in the context of our model (cf. Panel D of Figure 7).

Panel D 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 in Panel D are consistent with the joint hypothesis that our model is true and that the new RR technology was widely adopted around 1857.

5.3. Other Evidence

Three recent papers explicitly test some predictions of our model. Bharath and Viswanathan (2006) empirically analyze the model's risk implications at the firm level. They examine 252 brick-and-mortar firms that launched commercial websites (i.e., adopted the Internet technology as a way of doing business) in 1995–2004. The authors find that adopting the new technology is associated with an increase in firm risk, with differences between the early and late adopters: Firms that adopted the Internet before March 2000 (while stock prices were rising) experienced significant increases in idiosyncratic risk, whereas firms that adopted after March 2000 had significant increases in systematic risk. The authors conclude that their evidence provides strong support for our model.

Mazzucato and Tancioni (2006) analyze stock prices and patent-related measures of innovation in a sample of firms in the pharmaceutical and biotechnology industries between 1975 and 1999. They find that the firms' price-earnings ratios are positively related to innovation as well as to idiosyncratic risk, and argue that this evidence supports our model.

Hoberg and Phillips (2006) empirically examine the real and financial outcomes following industry booms in 1972–2004. They test the risk predictions of our model at the industry level. They find that market betas increase and idiosyncratic risk declines after booms, consistent with our model. The authors find strong support for our model in competitive industries but not in concentrated industries. It would be useful to extend our simple model to analyze the effect of product market competition theoretically.

In earlier work, Mazzucato (2002) studies the early phases of the life-cycles of the automobile and PC industries in the U.S. She finds that in both industries, stock prices were the most volatile when technological change was the most radical. She also finds idiosyncratic risk to be higher in the early stage of industry evolution, consistent with our model.

6. Conclusions

We offer 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 stock prices are observable only in hindsight – this pattern is unexpected while investors are uncertain whether the new technology would be widely adopted, but we observe it ex post when we focus on technologies that eventually led to technological revolutions. These "bubbles" should be most pronounced in revolutions characterized by high uncertainty and fast adoption. To formalize the intuition, we develop a general equilibrium model that features a real option decision and Bayesian learning about the average productivity of the new technology.

The model makes many empirical predictions. We find substantial support for these predictions in the evidence from 1830–1861 and 1992–2005 when the railroad and Internet technologies spread in the United States. In the context of our model, the empirical evidence is consistent with large-scale adoptions of railroads by the late 1850s and the Internet by 2002. A systematic empirical study of stock prices during all technological revolutions is beyond the scope of this paper, but it is a promising avenue for future research.

Future research can also test our model against alternatives that involve behavioral biases. To construct a fair horserace, it would be useful to develop a behavioral model of technological revolutions that produces testable predictions. Some predictions of our model, such as those involving market beta, are unlikely to follow from behavioral models, in which there is typically no role for systematic risk. Since we find empirically that systematic risk increased sharply during two prominent revolutions, it seems unlikely that behavioral biases can fully explain the observed stock price patterns. Such biases could certainly be part of the story, though, and quantifying their relative importance would be interesting.

Although we focus on stock prices, our model also has some implications for productivity. 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).

Our simple model has no 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.[21] Such an extension might have novel implications for investment while preserving the pricing implications of our model.

[21] 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. In contemporaneous work, DeMarzo, Kaniel, and Kremer (2006) develop a model in which a different mechanism – relative wealth concerns, arising due to limitations in the ability to trade future endowments and competition over future consumption – leads to overinvestment during technological revolutions.

Table 1

Parameters used in Simulations.

	$\overline{\rho}$	$\widehat{\psi}_{t^*}$	$\widehat{\sigma}_{t^*}$
	0.1217	0	0.04
ϕ	σ	$\sigma_{N,0}$	$\sigma_{N,1}$
0.3551	0.07	0.07	0.07
κ	$t^{**} - t^*$	T	γ
0.1	8	30	4

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

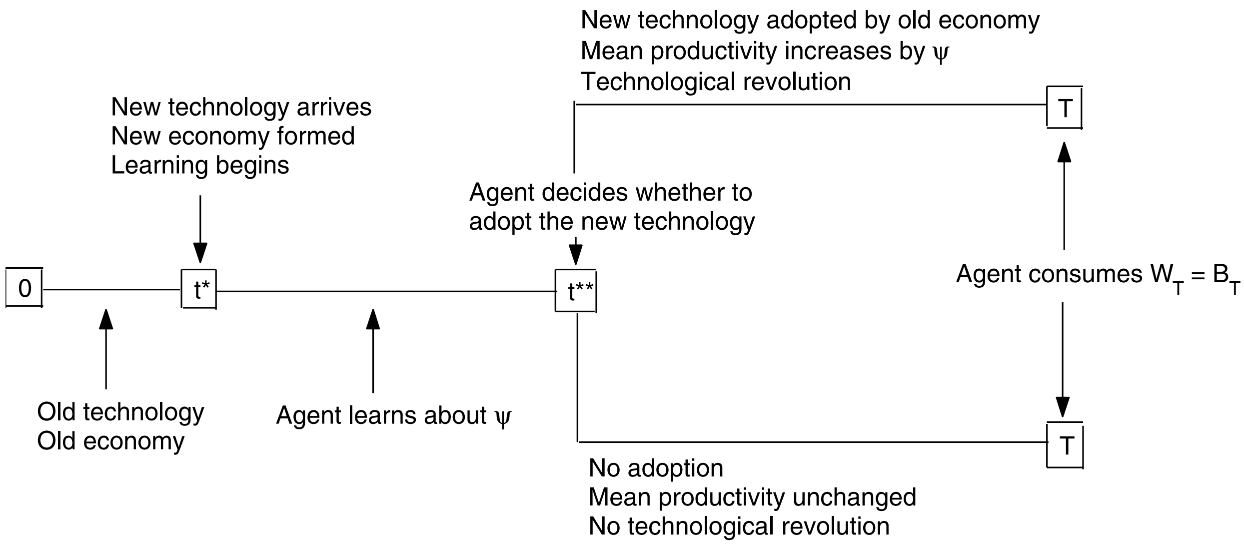

Figure 1. The Sequence of Events. In this chart, t^{**}, the time when the agent decides whether to adopt the new technology, is taken as given. We initially take t^{**} as given for the purpose of obtaining closed-form solutions for prices, but later we solve for the optimal time t^{**} to adopt the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

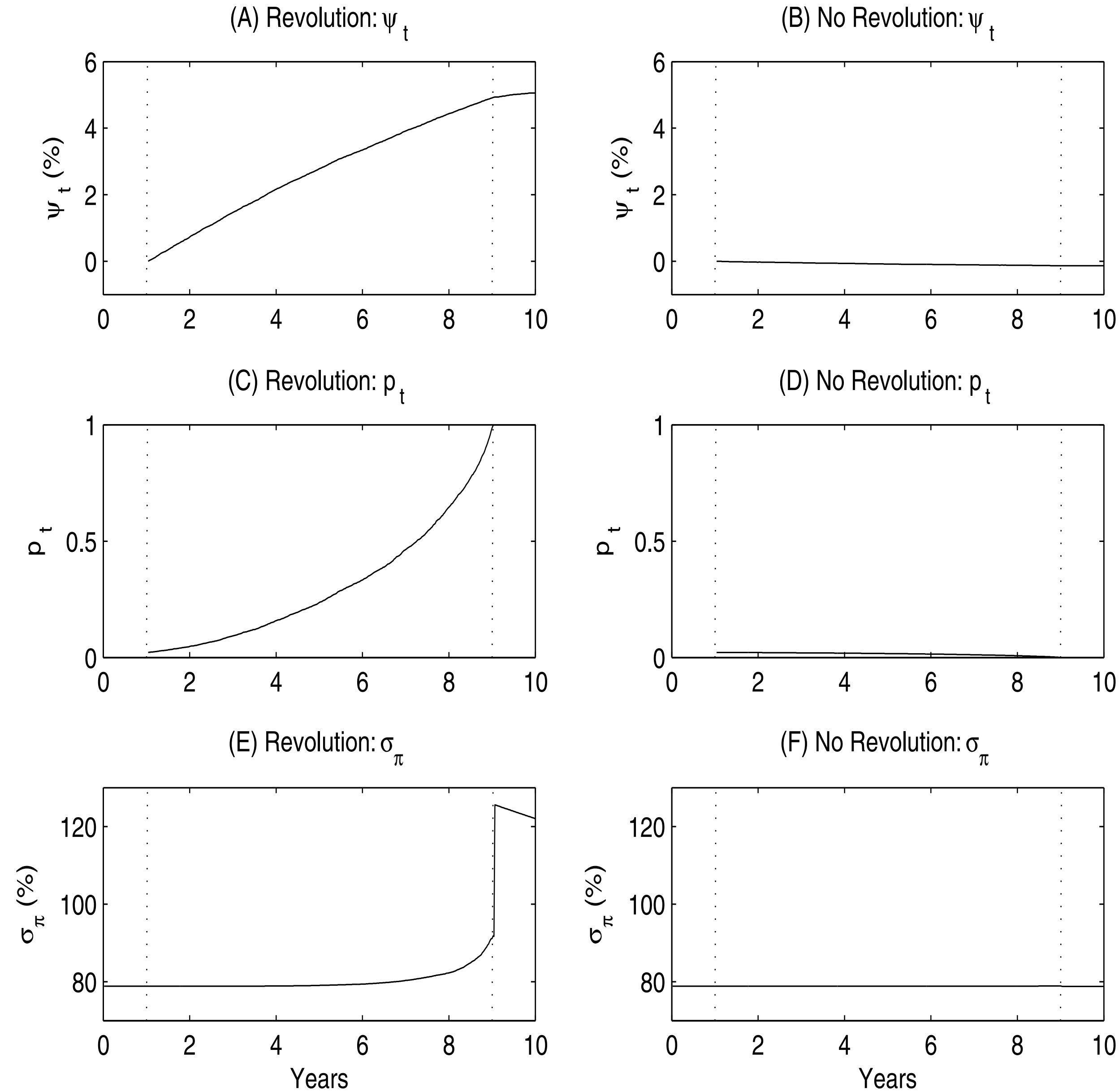

Figure 2. Average $\widehat{\psi}_t$, p_t, $\sigma_{\pi,t}$ in Simulations. The left panels plot the perceived productivity gain $\widehat{\psi}_t$ (Panel A), the adoption probability p_t (Panel C), and the volatility of the stochastic discount factor $\sigma_{\pi,t}$ (Panel E), averaged across all simulations in which the new technology was adopted at time t^{**} ($p_{t^{**}} = 1$). The right panels (B, D, and F) 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.

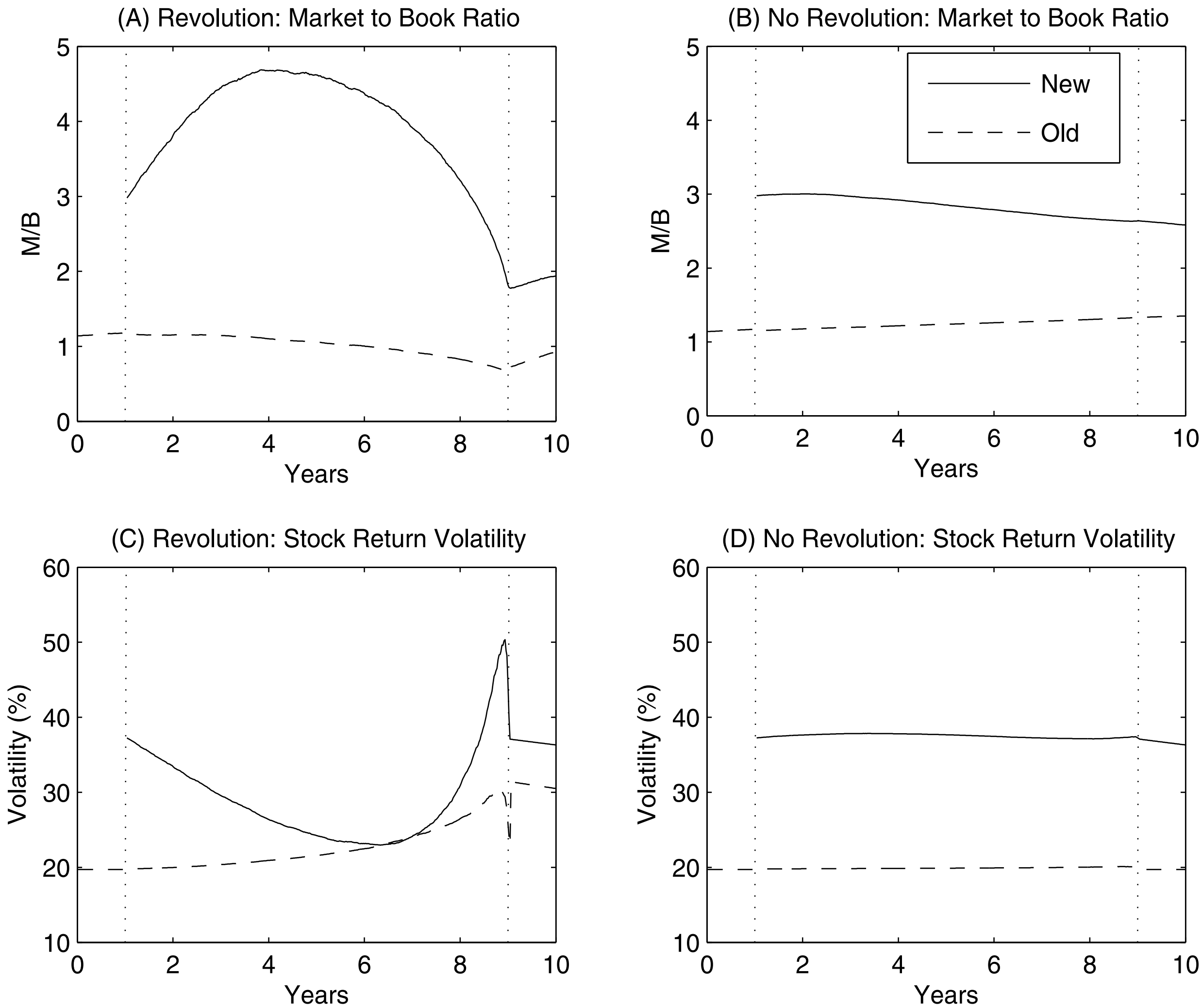

Figure 3. Average M/B and Volatility in Simulations. Panel A plots the path of the market-to-book ratio of the new economy (solid line) and old economy (dashed line) averaged across all simulations in which the new technology was adopted at time t^{**} ($p_{t^{**}} = 1$). Panel B is an equivalent of 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.

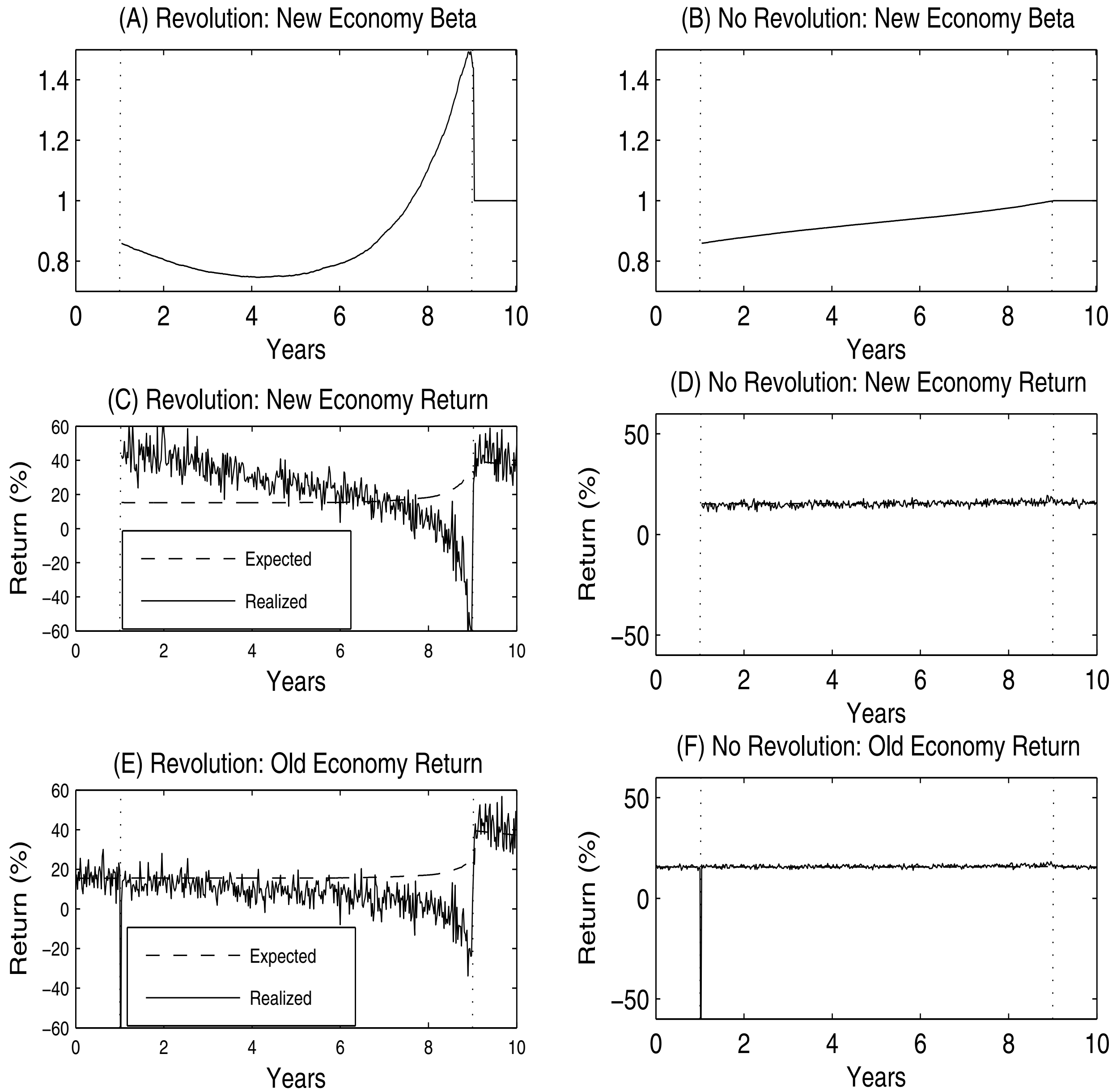

Figure 4. Beta and Average Stock Return in Simulations. The left panels plot the beta of the new economy (Panel A), the realized stock return (solid line) and the expected stock return (dashed line) for the new economy (Panel C) and the old economy (Panel E), averaged across all simulations in which the new technology was adopted at time t^{**} ($p_{t^{**}} = 1$). The right panels (B, D and F) 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.

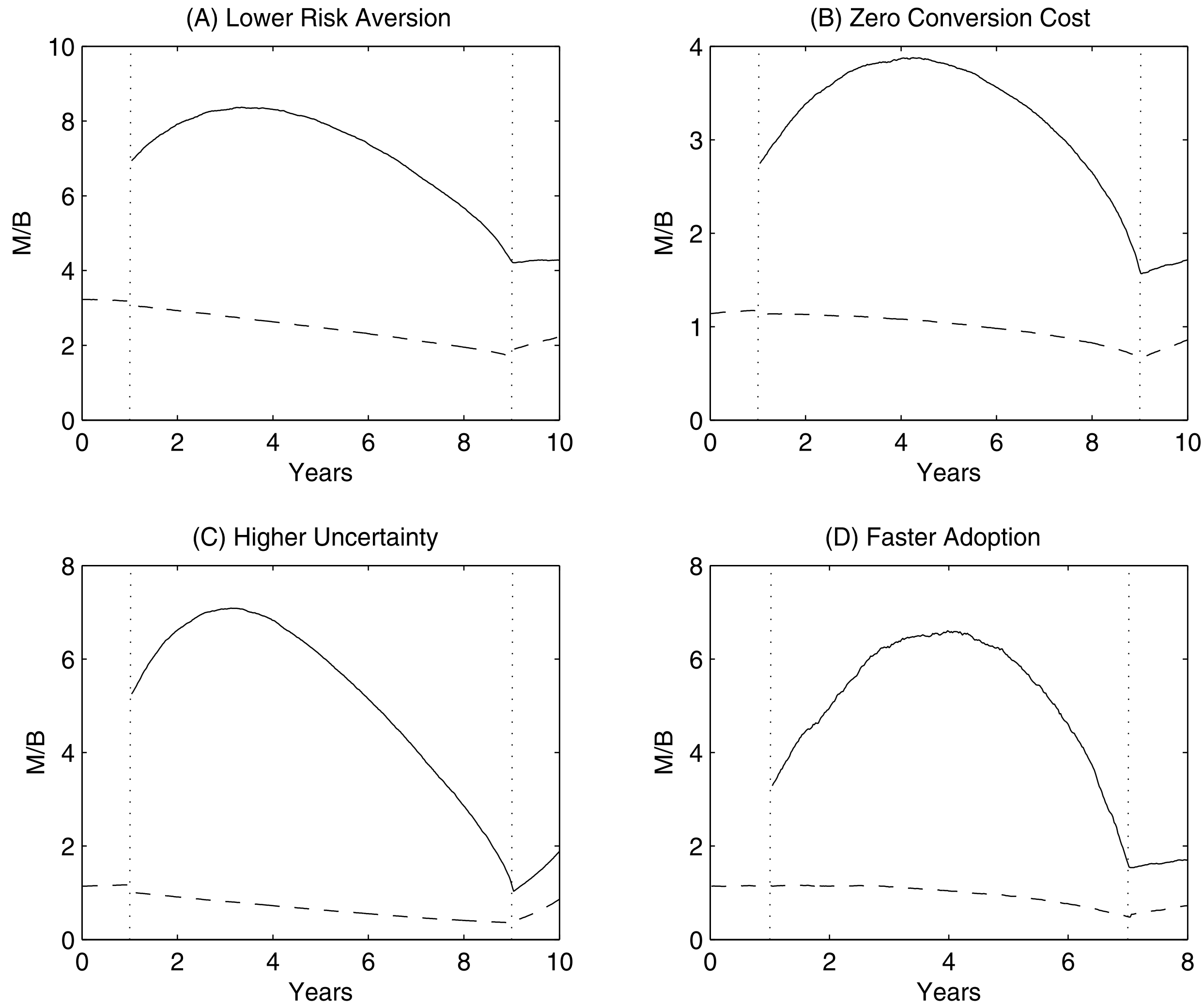

Figure 5. Average M/B in Simulated Revolutions: Sensitivity Analysis. All four panels plot the paths of the market-to-book ratio of the new economy (solid line) and old economy (dashed line) averaged across all simulations in which the new technology was adopted at time t^{**} ($p_{t^{**}} = 1$). All parameters are given in Table 1, except for one change that varies across the panels. In Panel A, the risk aversion $\gamma = 3$ instead of the benchmark case $\gamma = 4$. 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.

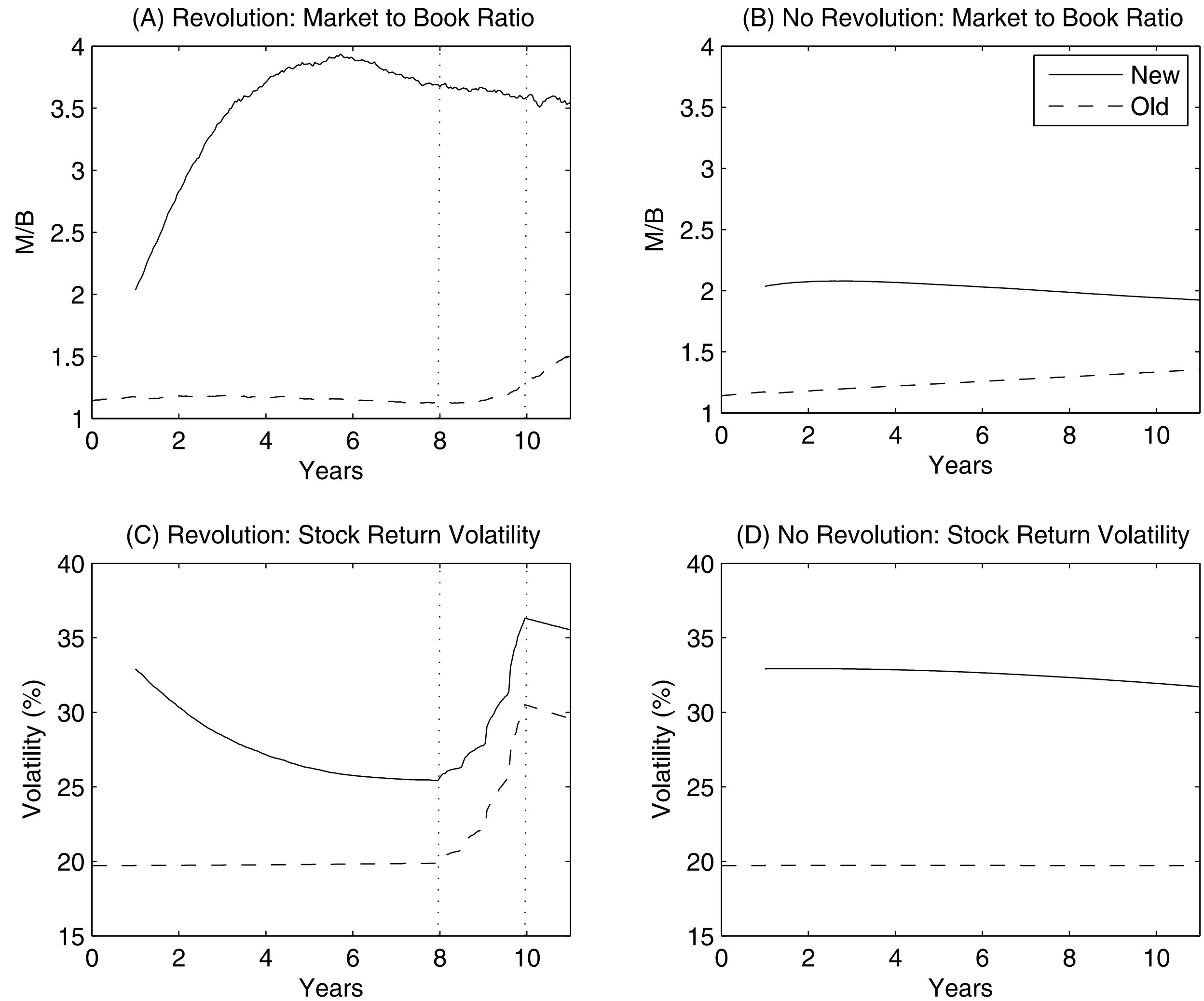

Figure 6. Average M/B and Volatility in Simulations with Optimal Adoption Time. Panel A plots the path of the market-to-book ratio of the new economy (solid line) and old economy (dashed line) averaged across all simulations in which the new technology was adopted at an optimally chosen time t^{**} between 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.

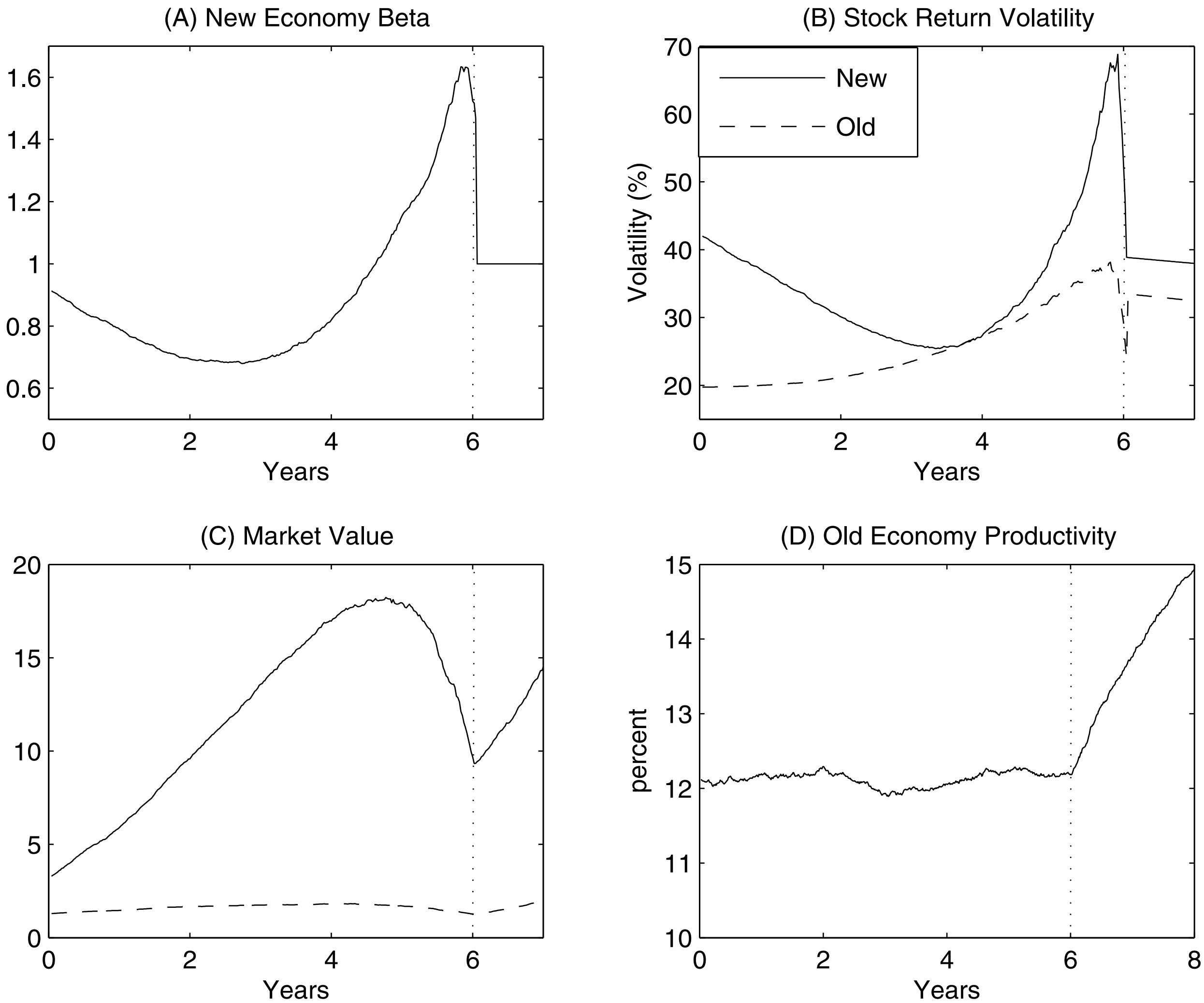

Figure 7. The Internet Revolution: Theory. This figure plots the model-implied dynamics of selected quantities in a revolution characterized by fast adoption of the new technology (such as the internet revolution). All quantities are averages computed across all simulations that led to a revolution (i.e., adoption at time $t^{**} = 6$. Panel A reports the market beta of the new economy. Panel B plots the market volatility of the new economy (solid) and old economy (dashed). Panel C plots the market values of the new economy (solid) and old economy (dashed). Panel D plots the productivity of the old economy.

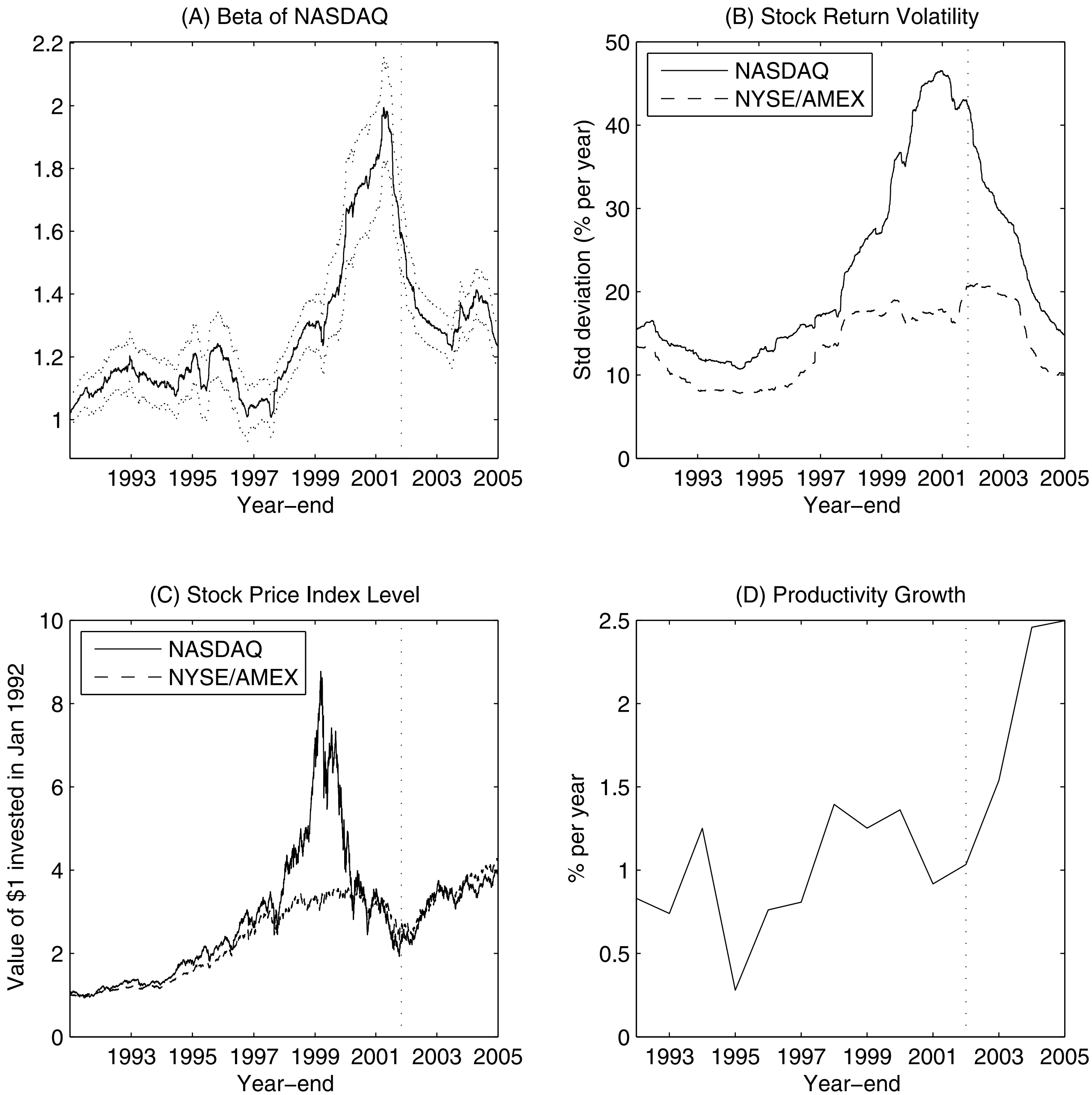

Figure 8. The Internet Revolution: Data. Panel A plots the market beta of the NASDAQ index, with a two-standard-error confidence band. The beta in day t is the slope coefficient from the regression of daily NASDAQ returns on NYSE/AMEX returns over the 500 trading days (i.e., about two years) immediately preceding day t (i.e., days $t-499$ through t). Panel B plots the standard deviations of returns on the NASDAQ and NYSE/AMEX indices, also computed from daily data over the 500 trading days immediately preceding day t. Panel C plots the value of $1 invested in January 1992 in the NASDAQ and NYSE/AMEX indices, with dividend reinvestment. Panel D plots a three-year moving average of multifactor productivity growth in the private business sector (in year t, we plot the average annual productivity growth in years $t-2$, $t-1$, and t). The vertical dotted line marks October 2002.

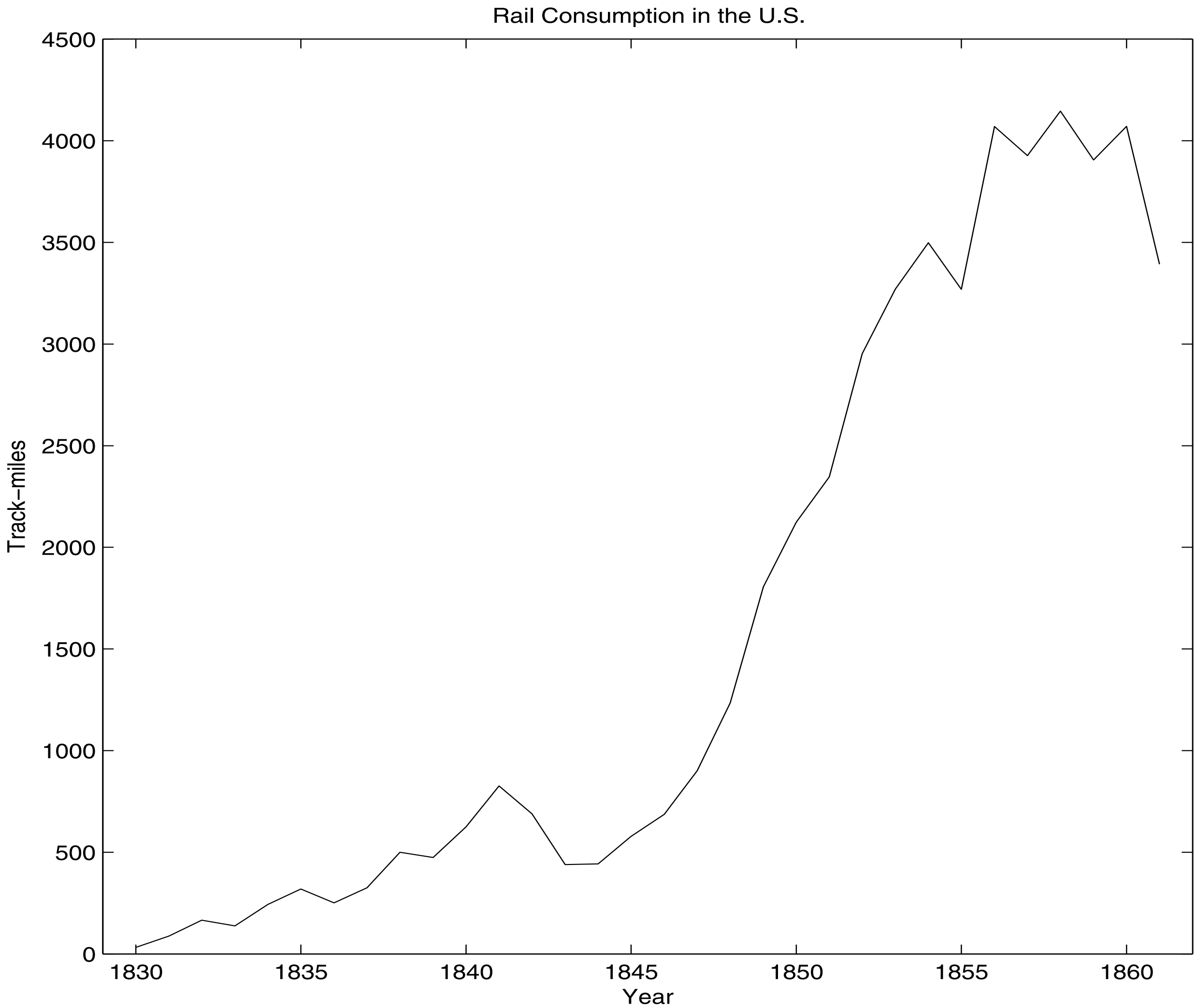

Figure 9. Total Rail Consumption in the United States. The figure plots the number of track-miles of rails laid each year in the U.S., as estimated by Fogel (1964, p.174). A track-mile of rails is defined as one half of the length of the rails in a mile of single track. The total includes rails used in the construction of new track as well as in the replacement of worn-out rails.

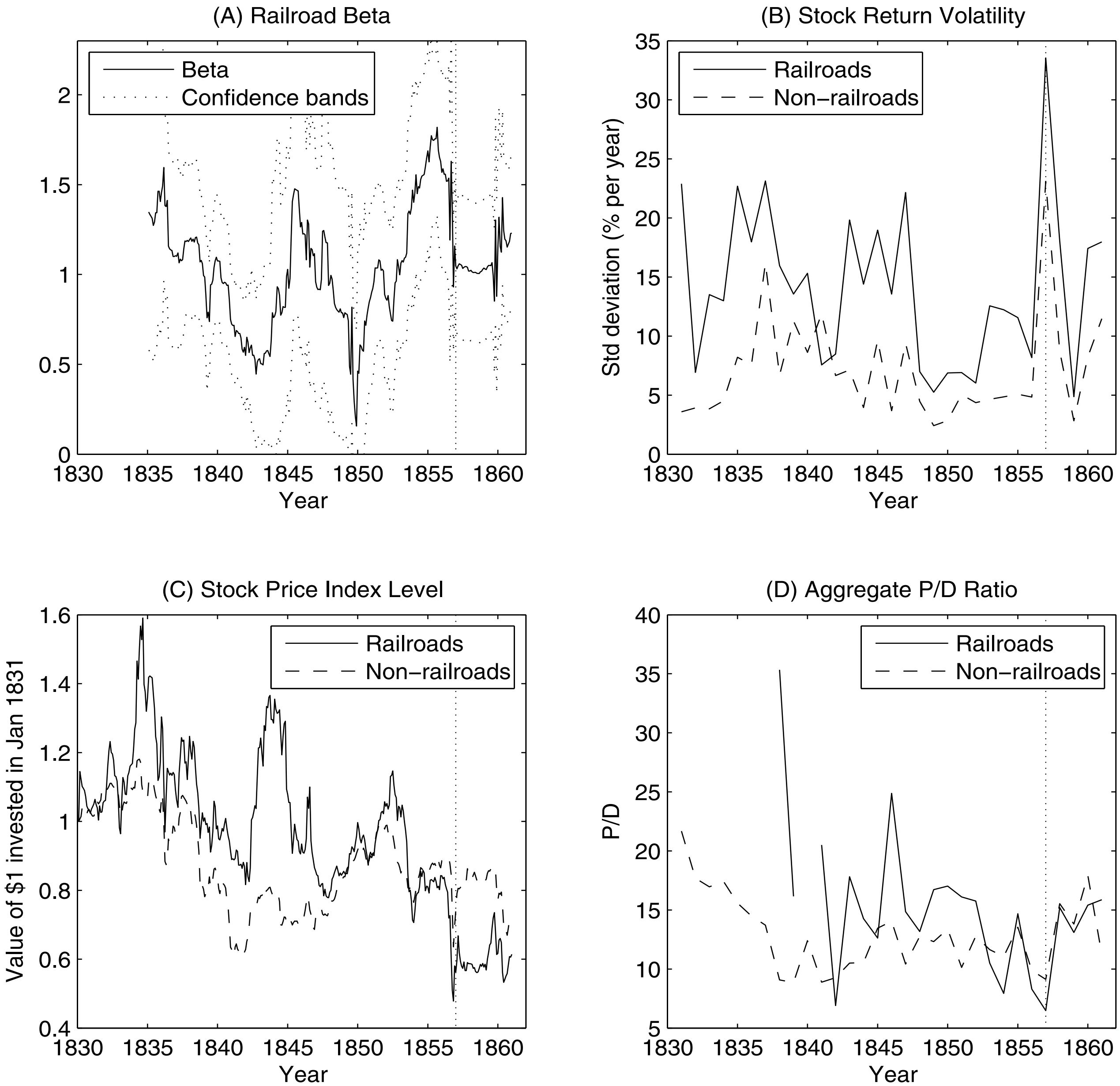

Figure 10. The Railroad Revolution: Data. Panel A plots the beta of the railroad index, with a two-standard-error confidence band. The beta in month t is the slope coefficient from the regression of the most recent 36 monthly railroad returns (in months $t-35$ through t) on non-railroad returns. The monthly railroad (non-railroad) index returns are computed as price-weighted averages of monthly capital gains across all railroad (non-railroad) stocks. Panel B 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 C plots the stock price index level, obtained by cumulating monthly capital gains to $1 invested in January 1831 in the railroad and non-railroad indices. Panel D plots the aggregate price-to-dividend ratio for the railroad and non-railroad 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. The vertical dotted line marks 1857.

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

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

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 "expectation errors" $(\widetilde{Z}_{0,t}, \widetilde{Z}_{1,t})$, which capture perceived innovations in ρ_t^N and ρ_t, follow

$$\begin{pmatrix} d\tilde{Z}_{0,t} \\ d\tilde{Z}_{1,t} \end{pmatrix} = \begin{pmatrix} \sigma & 0 \\ \sigma_{N,0} & \sigma_{N,1} \end{pmatrix}^{-1} \left(\begin{array}{c} d\rho_t \\ d\rho_t^N \end{array} - E_t\begin{bmatrix} d\rho_t \\ d\rho_t^N \end{bmatrix}\right). \tag{37}$$

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, $p_t \equiv \text{Prob}\left(\widehat{\psi}_{t^{**}} > \underline{\psi}|\widehat{\psi}_t\right) = 1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \widehat{\sigma}_t^2 - \widehat{\sigma}_{t^{**}}^2\right)$, where $\mathcal{N}(x; a, s^2) \equiv \int_{-\infty}^{x} (2\pi s^2)^{-1/2} \exp(y-a)^2/(2s^2)dy$ is the cdf of a normal distribution, $N(a, s^2)$.

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

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

$$\mathcal{V}\left(B_t, \rho_t, \widehat{\psi}_t, \widehat{\sigma}_t^2, t; T\right) = \frac{B_t^{1-\gamma}}{1-\gamma}\left\{(1-p_t)\,G_t^{no} + p_t G_t^{yes}\right\}, \tag{38}$$

where

$$G_t^{no} = e^{A_0(\tau)+(1-\gamma)A_1(\tau)\rho_t} \tag{39}$$

$$G_t^{yes} = G_t^{no}(1-\kappa)^{1-\gamma} R_t e^{(1-\gamma)A_2(\tau^{**})\widehat{\psi}_t+\frac{1}{2}(1-\gamma)^2 A_2(\tau^{**})^2\widehat{\sigma}_t^2} \tag{40}$$

and

$$R_t = \frac{1-\mathcal{N}\left(\underline{\psi};\widehat{\psi}_t+(1-\gamma)A_2(\tau^{**})\left(\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right),\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)}{1-\mathcal{N}\left(\underline{\psi};\widehat{\psi}_t,\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)} < 1. \tag{41}$$

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, $V(B_{t^*},\rho_{t^*},0,0,t;T^*) = B_{t^*}^{1-\gamma}/(1-\gamma)G_{t^*}^{no}$. Comparing this formula with $\mathcal{V}(B_{t^*},\rho_{t^*},0,\widehat{\sigma}_{t^*}^2,t^*;T)$ 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

$$\widetilde{G}_t^{yes} \equiv E\left[\left(\frac{B_T}{B_t}\right)^{-\gamma}|\widehat{\psi}_{t^{**}}\geq\underline{\psi}\right] = \widetilde{G}_t^{no}(1-\kappa)^{-\gamma}\widetilde{R}_t e^{-\gamma A_2(\tau^{**})\widehat{\psi}_t+\frac{1}{2}\gamma^2 A_2(\tau^{**})^2\widehat{\sigma}_t^2}, \tag{42}$$

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

$$\widetilde{R}_t = \frac{1-\mathcal{N}\left(\underline{\psi};\widehat{\psi}_t-\gamma A_2(\tau^{**})\left(\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right),\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)}{1-\mathcal{N}\left(\underline{\psi};\widehat{\psi}_t,\widehat{\sigma}_t^2-\widehat{\sigma}_{t^{**}}^2\right)} < 1 \tag{43}$$

$$\overline{A}_0(\tau) = -\gamma\overline{\rho}(\tau-A_1(\tau)) + \frac{\sigma^2}{2}\frac{\gamma^2}{\phi^2}\left\{\tau+\frac{1-e^{-2\phi\tau}}{2\phi}-2\frac{1-e^{-\phi\tau}}{\phi}\right\}. \tag{44}$$

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

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

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

$$S_{M,t} = \frac{-\frac{\partial p_t}{\partial \widehat{\psi}_t}G_t^{no} + \left((1-\gamma)A_2(\tau^{**}) + \frac{1}{\overline{p}_t}\frac{\partial \overline{p}_t}{\partial \widehat{\psi}_t}\right)G_t^{yes}}{(1-p_t)G_t^{no} + p_t G_t^{yes}}. \tag{46}$$

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

$$S^N_{M,t} = \frac{\left(A_2(\tau) + \frac{1}{p^N_{L,t}}\frac{\partial p^N_{L,t}}{\partial \widehat{\psi}}\right) K^{no}_t + \left((A_2(\tau) - \gamma A_2(\tau^{**})) + \frac{1}{p^N_{H,t}}\frac{\partial p^N_{H,t}}{\partial \widehat{\psi}}\right) K^{yes}_t}{(1-p_t)K^{no}_t + p_t K^{yes}_t}. \qquad (47)$$

Proof of Proposition 6. First, we rewrite the M/B ratio 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 $\widetilde{H}_t = \widetilde{G}^{yes}_t - \widetilde{G}^{no}_t$. Taking the derivative $\partial MB_t/\partial p_t$, we find that M/B increases in p_t if and only if $H_t \widetilde{G}^{no}_t > G^{no}_t \widetilde{H}_t$. Substituting the closed-form expressions, we obtain the condition $h_{old} > 0$, where

$$h_{old} = -\widetilde{\kappa} + A_2(\tau^{**})\widehat{\psi}_t + \frac{1}{2}(1-2\gamma) A_2(\tau^{**})^2 \widehat{\sigma}^2_t \qquad (48)$$

$$- \log\left(\frac{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2(\tau^{**})(\widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}), \widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right)}{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + (1-\gamma) A_2(\tau^{**})(\widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}), \widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right)}\right). \qquad (49)$$

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

$$MB^N_t = \frac{K_t R^N_L + p_t \bar{J}_t}{\widetilde{G}^{no}_t + p_t \widetilde{H}_t},$$

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

$$h_{new} = -\gamma A_2(\tau^{**}) A_2(\tau) \widehat{\sigma}^2_t - \log\left(\frac{\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t + \sigma^L_{y\widehat{\psi}}, \widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right)}{\mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t, \widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right)}\right) \qquad (50)$$

$$- \log\left(\frac{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2(\tau^{**})(\widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}), \widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right)}{1 - \mathcal{N}\left(\underline{\psi}; \widehat{\psi}_t - \gamma A_2(\tau^{**})(\widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}) + \sigma^L_{y\psi}, \widehat{\sigma}^2_t - \widehat{\sigma}^2_{t^{**}}\right)}\right). \qquad (51)$$

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,

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

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

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

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

$$\mathcal{V}\left(B_t,\rho_t,\widetilde{\psi}_t,\widehat{\sigma}^2_t,t;T\right)=B_t^{1-\gamma}e^{(1-\gamma)A_1(t)\rho_t}\mathcal{V}_2\left(\widehat{\psi}_t,t;T\right), \tag{52}$$

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

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

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

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

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

Proof: Since $\widehat{\sigma}^2_t$ 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

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

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 (52) with $\mathcal{V}_2$ satisfying the PDE and the boundary conditions given in Proposition 8.

REFERENCES

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

Atkeson, Andrew, and Patrick J. Kehoe, 2006, Modeling the Transition to a New Economy: Lessons from Two Technological Revolutions, *American Economic Review*, forthcoming.

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

Bharath, Sreedhar T., and Siva Viswanathan, 2006, Is the Internet bubble consistent with rationality?, Working paper, University of Michigan.

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.

Cochrane, W., 1979, *The Development of American Agriculture*, Minneapolis.

Craig, Lee A., and Thomas Weiss, 1993, Agricultural Productivity Growth During the Decade of the Civil War, *Journal of Economic History* 53, 527548.

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.

DeMarzo, Peter, Ron Kaniel, and Ilan Kremer, 2006, Technological Innovation and Real Investment Booms and Busts, *Journal of Financial Economics*, forthcoming.

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, *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.

Hoberg, Gerard, and Gordon Phillips, 2006, Real and financial industry booms and busts, Working paper, University of Maryland.

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 run-ups, Working paper, University of Michigan.

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

Mazzucato, Mariana, and Massimiliano Tancioni, 2006, Stock price volatility and patent citation dynamics, Working paper, Open University.

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* 81, 61–100.

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.