Asset Pricing Implications of Disruptive Technological Change

John Laitner and Dmitriy Stolyarov*

February 7, 2019

Abstract

This paper modifies an aggregative general equilibrium model with standard neoclassical features and isoelastic preferences to include random, low-frequency, disruptive technology shocks. We show that large-scale, disruptive shocks give rise to economic mechanisms that can produce a sizeable equity premium, a low risk-free rate, and stock returns that are both volatile and predictable. The model can thus provide a unifying perspective on a set of well-known asset pricing puzzles. In our numerical examples, we find a reassuring degree of agreement with the data. Our coefficients of relative risk aversion are between 1 and 1.5.

1 Introduction

This paper studies the asset pricing implications of disruptive technological change. Economic historians often suggest that technological progress proceeds in waves, with seminal inventions inaugurating new eras of accelerated growth. Major inventions can be highly disruptive, in some cases altering production methods and processes throughout the economy. In that vein, commentators identify revolutionary changes in sources of power and transportation during the 19th and early 20th centuries, followed by developments in microelectronics, information processing and communication, and the Internet.¹ Much recent attention focuses on potential future breakthroughs in

*John Laitner: University of Michigan, Department of Economics, 611 Tappan Ave, Ann Arbor, MI 48109-1220; Dmitriy Stolyarov: University of Michigan, Department of Economics, 611 Tappan Ave, Ann Arbor, MI 48109-1220; E-mail: stolyar@umich.edu. The authors thank seminar participants at the University of Michigan for insightful comments.

¹See, for example, Fogel (1964), David (1990), Mokyr (1990a, b) and Gordon (2016).
workplace automation. Given the prominence of disruptive technological change in such discussions, we examine whether embedding it into a workhorse economic growth model enhances the latter’s ability to explain asset market facts.

Questions persist about what types of macroeconomic uncertainty can drive risk premia in asset markets. For example, Cochrane (2008, p. 238) writes:

The challenge is to find the right measure of “bad times,” rises in the marginal utility of wealth, so that we can understand high average returns or low prices as compensation for assets’ tendency to pay off poorly in bad times.

We propose a new answer to this challenge. We present a model – at the intersection of macroeconomics and finance – with large-scale, randomly timed, punctuated technological advances. We show that our formulation is tractable for analysis and that it can offer new interpretations of linkages between financial markets and the real economy.

The model uses a number of standard neoclassical elements, including households that optimize over infinite horizons and a vintage-capital production sector (see Solow (1960) and Laitner and Stolyarov (2003)). Household preferences are isoelastic and our illustrative examples use a coefficient of relative risk aversion of 1-2. Two key assumptions are as follows. First, changes in the frontier TFP are large though infrequent. This captures economic historians’ notion of technological revolutions – or, equivalently, the idea that random arrivals of transformative general purpose technologies (GPT) drive economic growth (e.g., Helpman and Trajtenberg (1998)). Second, we assume that implementation of a major new technology requires technology-specific capital (e.g., Greenwood et al. (1997), Cummins and Violante (2002)).

**Preview of Results** We assess the model’s performance by confronting it with four well-studied aggregative asset-pricing facts: the sizable average equity premium, the low level of the risk-free rate, stock return predictability (e.g., Campbell and Shiller, 1988), and the high volatility of stock prices relative to dividends (Shiller, 1981). All four facts can be characterized as “puzzles” because they appear inconsistent with standard theories in macroeconomics and/or baseline hypotheses in finance.

In our model, the mechanisms governing asset-pricing phenomena depend upon the reaction of the stock market to a major technological innovation and the subsequent diffusion of the new technology. Importantly, disruptive technological change can cause large declines in stock prices. Newer vintages of capital provide more capital services (i.e., efficiency units) per dollar of investment. When a major new technology arrives, existing stock prices precipitously drop, in step with the cost of producing capital services. At the same time, the availability of highly productive new capital creates exceptional opportunities for investment. Household saving decisions determine

---

2 See, for example, Autor (2015), Arntz et al. (2017).
the economy’s response, thus controlling the speed of diffusion of the new technology. Household preferences for consumption smoothing make diffusion gradual.

An asset carries a risk premium if its returns are low when the marginal utility of wealth is high. In our model, as a new technology arrives, existing assets suffer an abrupt capital loss. Hence, realized common stock returns are low. At the same moment, new investment opportunities lead households to divert, temporarily, a larger share of their income to saving—i.e., to purchasing new assets. Consumption falls, therefore, exactly as stock returns plummet, creating a negative correlation between the marginal utility of wealth and equity returns. The risk-free rate level is affected as well. Specifically, the ever-present possibility of a technological revolution, with its accompanying exceptional investment opportunities, makes bonds, which preserve an investor’s principal during the lead-in to a new GPT, an attractive portfolio asset.

In the conventional view, stock prices respond to news about dividends. In our model, however, a GPT shock affects both investment-good supply prices and future dividends. In fact, the shock moves them in opposite directions. Investment-good supply prices set the value of equity shares; adjustments of the post-shock discount rate discount rate equate share prices to the value of future dividends.

The model can then explain both stock return predictability and the “volatility puzzle.” Between the arrivals of new GPTs, the economy converges to the steady-state capital stock corresponding to full adoption of the (current) frontier technology. Realized stock returns decline along the convergence trajectory as the marginal product of capital works its way down. Because convergence is gradual—and disruptive innovations are rare—the current price-to-earnings ratio may forecast future returns quite well, on average. On the other hand, while aggregate dividends reflect the slow moving average marginal product of capital, existing shares of common stock can incur sharp capital losses when a new GPT reduces the supply price of efficiency units. Thus, the volatility of asset prices can be high relative to that of dividends.

In the end, the model can rationalize four observations on aggregate asset prices that receive substantial attention in the literature. The model points to waves of technological change as a unified explanation for behavior in both real and financial markets. The new framework decouples low stock returns from contemporaneous movements in output, as disruptive technological change is bad news for existing assets but good news for productivity growth; thus, the model offers new insights into the nature of the “bad times” that drive the risk premium.

**Related Literature** The literature on the equity risk premium is vast, reflecting the subject’s central importance for both macroeconomics and finance. The inability of standard real business cycle frameworks to match both the equity premium and the risk-free rate is well-documented, dating back to the influential work of Mehra and Prescott (1985). In the years since, macroeconomists

---

3See also Kocherlakota (1990, 1996) and Jermann (1998).
have explored generalized preferences (e.g., Weil (1989), Campbell and Cochrane, 1999), inflexible factor markets (e.g., Boldrin et al., 2001), transaction costs (e.g., Mehra and Prescott, 2008), and limited stock market participation (e.g., Guvenen, 2009) — individually, or in combination — as ways to reconcile macroeconomic theory and asset market facts. The present paper takes the different approach: with the exception of limited stock market participation, we maintain standard assumptions — but focus on the role of disruptive technology shocks.

Our work complements a strand of literature that considers large shocks as a potential explanation for the equity premium puzzle. Rietz (1988) proposed the hypothesis that the possibility of rare, adverse events causing large declines in consumption — so-called “rare disasters” — might significantly contribute to the observed equity premium. Work by Barro (2006) and Barro and Ursua (2008) sought to provide empirical support by calibrating the probability and size of rare disasters from a large body of cross-country time series data. Gourio (2008) adds recoveries to the Rietz-Barro model to account for return predictability. Gabaix (2012) extends the Rietz-Barro framework to rare disasters of variable size.

Our work differs from the rare disaster literature with respect to the source of shocks and the modeling approach. Empirical assessments of Rietz’s hypothesis focus on adverse events: wars, crises, natural cataclysms, etc., and take the processes for dividends and consumption to be exogenous. Our paper, by contrast, emphasizes large shocks associated with disruptive technological change. The model has both production and capital accumulation decisions. The model’s mechanisms do not rely upon large fluctuations of output to generate co-movement of consumption and stock returns.

A related literature seeks to understand links between real business cycles and asset markets (e.g., Jermann (1998), Tallarini (2000), Boldrin et al., 2001). Within this literature, our work complements analyses of real business cycle models with investment-specific technological change (ISTC) — e.g., Christiano and Fisher (2003), Papanikolaou (2011), and Kogan and Papanikolaou (2013).

Papanikolaou (2011) is, perhaps, the most closely related to ours in that it features a similar mechanism for co-movement of the rate of ISTC and the marginal utility of wealth. Papanikolaou’s calibrations attribute the equity risk premium mostly to the volatility of the ISTC rate. Our paper differs in both modeling and focus. Papanikolaou (2011) develops a two-sector real business cycle model where the rate of ISTC changes in small increments. He focuses on the mapping between ISTC and the cross-section of stock returns. Our framework, by comparison, assumes low-frequency investment-specific technology shocks and studies the link between technology diffusion and time-series asset pricing facts. We employ a minimal set of parameters and assumptions and study ISTC

---

4 For comprehensive literature surveys, see Kocherlakota (1996) and Mehra (2008).
5 See also Nakamura et al. (2013) for updated calibrations.
in isolation from other sources of macroeconomic risk.

Our work extends the theoretical literature that seeks to characterize macroeconomic dynamics at the time of technological revolutions and to understand the links between technological change and large stock market fluctuations. Greenwood and Jovanovic (1999) and Manuelli (2000) focus on the behavior of macroeconomic variables and the stock market, given the anticipation of a major technological change. Pastor and Veronesi (2009) develop a macroeconomic model in which gradual learning about the new technology’s potential drives a boom-bust pattern in the stock market. The papers just mentioned consider one-time technological events, whereas the present paper models recurring, random episodes of disruptive technological change. This paper shares the production side of Laitner and Stolyarov (2003), though it adds consumption-saving decisions. The new framework thus becomes suitable for studying risk premia and asset-pricing phenomena related to technology diffusion.

The organization of this paper is as follows. Section 2 describes the model, Section 3 presents the solution method and characterizes equilibrium, Section 4 suggests calibrations, and Section 5 provides quantitative assessment of the model and the mechanisms driving the asset-pricing puzzles. Section 6 concludes.

2 Model

Our framework of analysis is a discrete-time, vintage capital model with recurring, large technology shocks and a production side similar to Laitner and Stolyarov (2003) and Solow (1960). A disruptive technological change introduces a higher-TFP production function that requires a new input, that is, a new vintage of capital. As new investment adds to the stock of the new vintage, the frontier technology diffuses, and aggregate productivity grows. The diffusion path is interrupted by the next disruptive technological change, at which point aggregate investment switches to the newest vintage, and another diffusion phase starts. The household side of the model consists of two types of agents and features limited participation in asset markets.

In Section 2.1, we derive the aggregate production function, factor prices, market value of capital stock, and the time paths for the prices of existing vintages. Section 2.2 describes the behavior of the two types of households.

2.1 Output, factor prices and the value of capital

Technological progress General purpose technologies are indexed by \( j \), an integer, with newer technologies having a bigger \( j \). At time \( t \), the economy has a total of \( J_t \) distinct technologies, where \( J_t \) is a Poisson random variable. Each time a Poisson event occurs, a new technology appears
and \( J \) increases by 1. The (exogenous) hazard rate is \( \lambda \), so that the average interval between new technologies is \( 1/\lambda \). Technologies are ranked with respect to their TFP level. GPT \( j \) has a distinct TFP level \( \theta_0 \cdot [\theta]^j, \theta > 1 \). Without loss of generality, set \( J_0 = 0 \).

**Vintage capital** Each technology \( j \) has a separate production function that requires a distinct type of capital, \( K_j \), that we call vintage \( j \). We have a one sector economy where output can be produced with multiple types of capital, \( \{K_j\}_{j \leq J} \), specific to general purpose technologies. GPTs themselves are not privately owned, but technology-specific investments that implement and embody GPTs are assumed to be private, rival goods. Output (the numeraire) is homogeneously divisible into consumption and investment. Investment expenditure \( I_t \) can be transformed one-to-one into capital of any vintage. Let \( I_{jt} \) be time-\( t \) investment in capital of vintage \( j \leq J_t \). At time \( t \), only vintage \( J_t \) capital is built because it has the highest resale value (see Proposition 1 below); so, \( I_t = I_{J_t} \). We assume that investment is irreversible, \( I_{jt} \geq 0 \). Thus, old vintages cannot be transformed into new. Every vintage experiences wear-and-tear depreciation at rate \( \delta \).

**Labor supply** The economy is populated by two types of agents. Both inelastically supply labor. Let \( l \) denote type-I labor supply, and \( \bar{l} \) denote type-II labor supply. Labor input is not vintage-specific, and it is a Cobb-Douglas composite of the two types of labor,

\[
L = l^\beta \cdot \bar{l}^{1-\beta}, \beta \in (0, 1).
\]

Without loss of generality, the total amount of each type of labor is normalized to 1.\(^{6}\) In what follows, type-I agents will own all the equities, and type II agents will not participate in the stock market.

**Production function** The aggregate production function is the sum of outputs across different vintages, with the output-maximizing allocation of labor:

\[
Y_t = \max_{l_{jt}, \bar{l}_{jt}} \sum_{j \leq J_t} \theta_0 \cdot [\theta]^j [K_{jt}]^\alpha [L_{jt}]^{1-\alpha}, \alpha \in (0, 1),
\]

s. t. \( \sum_{j \leq J_t} l_{jt} \leq 1, \sum_{j \leq J_t} \bar{l}_{jt} \leq 1, l_{jt} \geq 0, \bar{l}_{jt} \geq 0, L_{jt} = l_{jt}^\beta \cdot \bar{l}_{jt}^{1-\beta} \).

**Value of capital** Let \( P_{jt} \) denote the resale value of one unit of \( K_j \) at date \( t \), when the newest vintage is \( J = J_t \). The capital of the latest vintage has price \( P_{jt} = 1 \), equal to the marginal cost of transforming output into capital. The resale prices \( P_{jt} \) obtain from the no-arbitrage condition with respect to owning capital of different vintages. The condition states that each dollar of capital

\(^{6}\)With this normalization, \( Y_t \) in (1) is interpreted as output per worker, whereas the magnitude of \( \theta_0 \) registers our choice of units for labor supply.
should produce an equal amount of output per worker, regardless of its vintage. The relative price of vintages then equals the relative amount of capital services (efficiency units) they provide. We have

**Proposition 1:**

\[ P_{jt} = [b]^{j-j}, \quad b \equiv [\theta]^{-1/\alpha} < 1. \]

**Proof:** See Appendix.

Proposition 1 reveals the relative efficiency of different vintages, with capital vintage \( j \) providing \([\theta]^{1/\alpha} = 1/b \) times more efficiency units than vintage \( j - 1 \).

The value of the aggregate capital stock at date \( t \) is

\[ K_t = \sum_{j \leq J_t} P_{jt} \cdot K_{jt}. \]  

Expression (2) shows that the value of capital equals the aggregate quantity of vintage-\( J \) efficiency units. A new technology arrival reduces the value of existing capital by a factor \( b \). Intuitively, the newest vintage provides \( 1/b \) times more capital services per dollar of investment than the previous frontier, so the price of capital services drops by factor \( b \) economy-wide.

We now solve (1) for the aggregate production function. The following proposition parallels analysis in Solow [1960] and Laitner and Stolyarov [2003].

**Proposition 2:** Let \( \mathcal{Y}_t \) be as in (1) and \( K_t \) be as in (2). Then

\[ \mathcal{Y}_t = \theta_0 \cdot [\theta]^{J_t}[K_t]^{\alpha}. \]  

The net marginal revenue product per dollar of physical capital, of any vintage, is

\[ \mathcal{M}(K_t, J_t) = \alpha \frac{\mathcal{Y}_t}{K_t} - \delta, \]  

and, labor income of type-I agents is

\[ \mathcal{W}(K_t, J_t) = (1 - \alpha)\mathcal{Y}_t. \]

**Proof:** See Appendix.

Proposition 2 provides several key simplifications. On the one hand, we do not need to keep track of quantities \( K_{jt} \) separately. On the other hand, given inelastic labor supplies, there are just two aggregate state variables in the production sector: \( K_t \) and \( J_t \).
Detrending

Before formulating the household decision problem and defining equilibrium, it is convenient to make a change of variables that removes the growth trend. Let \( Z = \theta^{1-\alpha} \) denote the growth factor for capital, output and consumption associated with a TFP step \( \theta \). Define aggregate detrended variables as follows:

\[
K_t = \frac{K_t}{Z^{j_t}}, \quad Y_t = \frac{Y_t}{Z^{j_t}}, \quad w_t = \frac{W(K_t, J_t)}{Z^{j_t}},
\]

where

\[
w_t = w(K_t) = (1-\alpha)\beta \theta_0 K_t^\alpha
\]

is the detrended wage. In addition, let \( m_t \) (\( M_t \)) denote the net (gross) marginal revenue product per dollar of physical capital expressed through the detrended capital stock:

\[
M_t = M(K_t) = \alpha \theta_0 K_t^{\alpha-1}, \quad m_t = m(K_t) = M(K_t) - \delta.
\]

We next describe the behavior of households.

2.2 Households

The economy has two types of households, “wealthy” type-I and “middle class” type-II.\(^7\) A type-I household is infinitely lived, it receives labor and capital income, and it makes consumption-saving choices. Our analysis excludes residential fixed assets from capital and housing services from output. We assume that only type-I households participate in the stock market and finance the economy’s non-residential capital stock. Type-I households receive all of the economy’s capital income and fraction \( \beta \) of aggregate earnings, with their overall share of aggregate income equal to \( \alpha + \beta \left(1 - \alpha\right) = \eta \). Type-II households receive just labor income, in the amount of \( (1-\beta)(1-\alpha)Y_t = (1-\eta)Y_t \). In effect, type-II households just consume their labor earnings.\(^8\) The role of the limited participation assumption is to discipline the model’s quantitative predictions with respect to fluctuations of aggregate consumption (see Table 1); the paper’s qualitative results and analytical characterizations apply for any \( \beta \in [0, 1] \).

The remainder of this section focuses on the behavior of type-I households. It is convenient to first define a new variable \( z_t = J_t - J_{t-1} \in \{0, 1\} \) that registers whether a new technology arrives at date \( t \) or not. As above, set \( J_0 = 0 \). Then \( J_t = \sum_{s=1}^{t} z_s \). The history of technology shocks up to date \( t \) is denoted \( z^t = \{z_s\}_{s=1}^{t} \), and the set of all possible histories up to date \( t \) is denoted \( Z^t \), with

\(^7\)Our calibrations will associate the type-I household with those in the top 5 percent of the US wealth distribution.

\(^8\)We could allow type-II households to finance their own residences. For simplicity, however, we exclude housing from the model.
probability of history $z^t$ given by

$$p(z^t) = \lambda^J (1 - \lambda)^{t-J_t}. \quad (8)$$

Our timing convention in (6)-(7) is to define date-$t$ variables at the moment the current technology level $J_t$ is known.

We use the following notation for the consumption and wealth of type-I households, with capital letters denoting aggregate variables, and lower case letters denoting household-level variables: $C_t$ ($c_t$) the consumption level; $X_t = C_t/Z^{J_t}$ ($x_t$) detrended consumption; $A_t$ ($a_t$) detrended beginning of period-$t$ wealth before $z_t$ is realized; and, $K_t$ ($k_t$) the detrended market value of the capital stock after $z_t$ is realized. Below, we refer to a representative type-I household simply as a “household”.

Consider the household budget constraint and the law of motion for wealth. The current aggregate state $K_t$ determines factor prices $m(K_t)$ and $w(K_t)$. Let

$$A(k_t, K_t) \equiv k_t + m(K_t) k_t + w(K_t)$$
denote the initial period-$t$ resources of a household with current wealth $k_t$, when aggregate capital is $K_t$. Household wealth carried into next period, $a_{t+1}$, equals the difference between initial resources and consumption,

$$a_{t+1} = A(k_t, K_t) - x_t.$$  

With our timing convention, $a_{t+1}$ is determined before $z_{t+1}$ is known. At the beginning of period $t+1$, $z_{t+1}$ is realized. If there is no technology shock, $J$ and the price of capital are unchanged from the previous period, so $k_{t+1} = a_{t+1}$. If, however, a new technology arrives in period $t+1$, the aggregate capital stock is re-valued by a factor $b < 1$, and, in addition, the trend variable, $Z^J$, rises by a factor $Z > 1$. Accordingly, when $z_{t+1} = 1$, the realized household wealth for period $t+1$ is $k_{t+1} = ba_{t+1}/Z$. Thus, the law of motion for $k$ is

$$k_{t+1} = a_{t+1} \omega(z_{t+1}) = (A(k_t, K_t) - x_t) \omega(z_{t+1}), \quad (9)$$

where

$$\omega(z) \equiv \frac{b}{Z} z + (1 - z), \quad z \in \{0, 1\}.$$  

In a similar fashion, the law of motion for $a_t$ is

$$a_{t+1} = A(a_t \omega(z_t), K_t) - x_t.$$  

To formulate the household decision problem, we need to specify how households form their
expectations about factor prices $m(K)$ and $w(K)$. In other words, we need a law of motion for the aggregate state. Suppose that households have a common belief about the aggregate consumption decision, $X = \Phi(K)$. The law of motion for the aggregate state is then the aggregate version of (9):

$$K_{t+1} = \Gamma(K_t, z_{t+1}; \Phi) \equiv (\mathcal{A}(K_t, K_t) - \Phi(K_t)) \omega(z_{t+1}).$$  \hspace{1cm} (10)

The detrended capital stock trajectory, after any history $z^t$, can be constructed recursively from this law of motion.

Household preferences over consumption are

$$u(c) = \frac{c^\gamma}{\gamma}, \gamma < 1, \gamma \neq 0,$$

and the subjective discount factor, $\rho$, which we restrict to ensure that the household expected lifetime utility is finite,

$$\rho (1 - \lambda) + \rho \lambda Z^\gamma < 1. \hspace{1cm} (11)$$

The household problem involves choosing desired wealth for next period $a_{t+1} = a[z^t]$ after any history $z^t$, prior to knowing $z_{t+1}$. The household maximizes its expected lifetime utility, with the expectation taken over histories in $\mathcal{Z}^t$, and with the probability of history $z^t$ as in (8). The feasible set for the control variable $a[z^t]$ is the set $\mathcal{K}$ of aggregate states visited by rational expectation equilibrium trajectories. This set is an equilibrium object that Proposition 3 below characterizes. The period-$t$ household payoff function is

$$U(a_t, a_{t+1}, z^t) = u(Z^h : [A(a_t \omega(z_t), K[z^t]) - a_{t+1}]),$$

with $Z^h$ multiplying detrended variables to recover the household consumption level $c_t$. The household’s overall maximization problem, setting $a[z^{-1}] \omega(z_0) = k_0$, is

$$V(K_0, k_0; \Phi) = \sup_{\{a[z^t]\}_{t=0}^\infty} \sum_{t=0}^\infty \sum_{z^t \in \mathcal{Z}^t} p(z^t) \rho^t U(a[z^{-1}], a[z^t], z^t)$$

s.t. $a[z^t] \in \mathcal{K}$, $k_0 = K_0 \in \mathbb{K}$.  \hspace{1cm} (13)

We next proceed to the definition and characterization of equilibrium.

## 3 Equilibrium

A rational expectations equilibrium in this paper has three requirements: (1) households solve utility maximization problem (13) given common beliefs $\Phi(\cdot)$ and aggregate state forecast $K[z^{t+1}]$
constructed from (10); (2) beliefs are consistent with the household consumption policy function solving (13), so that the aggregate state trajectory coincides with its forecast; and, (3) the aggregate state trajectory has non-negative gross investment, so that investment irreversibility constraint is never binding for vintage $J_t$. The formal definition of equilibrium is as follows.

**Definition** An equilibrium is a feasible set $\mathbb{K}$; a list of sequences $\{\alpha[z^{t-1}], k[z^t], K[z^t]\}_{t=0}^{\infty}$ with $k[z^t] = a[z^{t-1}] \omega(z_t)$; and a forecast function $\Phi(K)$ satisfying the following conditions:

1. The policy function $\alpha[z^t]$ solves household problem (13) given forecast function $\Phi$ and the law of motion for $K[z^t]$ defined in (10);
2. The wealth trajectory chosen by the household coincides with the aggregate state trajectory

$$k[z^t] = K[z^t] \in \mathbb{K}, \text{ all } t \text{ and } z^t \in \mathbb{Z}^t; \tag{14}$$

3. Investment is non-negative along the equilibrium path,

$$\alpha[z^t] \geq (1 - \delta) K[z^t], \text{ all } t \text{ and } z^t \in \mathbb{Z}^t.$$

Consistency condition (14) is equivalent to

$$\Phi(K[z^t]) = x[z^t], \text{ all } t \text{ and } z^t \in \mathbb{Z}^t, \tag{15}$$

a requirement that detrended household consumption coincide with its aggregate forecast along the rational expectations path – compare (9) to (10).

### 3.1 Solution methodology

This section explains how we solve for a rational expectations equilibrium. The method involves constructing an aggregate consumption function $\Phi(K) = X$ that satisfies rational expectations condition (15). We solve for a growth trajectory satisfying (15) for a succession of TFP processes, each limiting the number of future GPT arrivals that households expect. For the case $n = 0$, there are no future GPT arrivals, the household problem is non-stochastic, and solving for the aggregate consumption function $\Phi^0(.)$ is straightforward. For $n = 1$, households expect a single future GPT arrival (with hazard $\lambda$). We show that we can construct $\Phi^1(.)$ consistent with rational expectations prior to the new GPT, using $\Phi^0(.)$ from the previous step in the continuation problem. Similarly, at step $n$, $\Phi^n(.)$ is defined recursively from $\Phi^{n-1}(.)$. As the number of future GPT arrivals, $n$, becomes larger, we anticipate that the difference between $\Phi^n(.)$ and $\Phi^{n-1}(.)$ will become ever smaller. Proposition 3 in Section 3.2 shows that if, in fact, the sequence $\Phi^n(.)$ converges uniformly, then its limit, $\Phi^*(.)$, defines a rational expectations equilibrium.
Figure 1: Phase diagram with $\lambda = 0$, $z_t = 0$.

Start with the case $n = 0$, where households believe that no technological change will ever occur. The version of the model with no technological change has $\lambda = 0$ and $z_t = 0$ for all $t$. That is a familiar framework where equilibrium is described by Euler and budget equations

$$u'(X_t) = \rho \cdot [1 + m(A_{t+1})] \cdot u'(X_{t+1}),$$

$$A_{t+1} = K_t + m(K_t) \cdot K_t + w(K_t) - X_t.$$ (16)

The Euler equation comes from differentiating (13) with respect to $a_{t+1}$. The budget equation (17) is the law of motion (9). If we set $a_{t+1} = A_{t+1}$ in (16) and $k_t = K_t$, $z_{t+1} = 0$, and $K_{t+1} = A_{t+1}$ in (17), the solution to (16)-(17) will be consistent with rational expectations.

Suppose the phase diagram is as in Fig 1, with a saddlepath solution. Suppose for any starting $K_0 = k_0 = K \in (0, \bar{K}]$, the saddlepath is (i) non-decreasing, (ii) continuous, and (iii) manifests direct motion toward the stationary point (at $\bar{K}$). From the height of the saddlepath above each starting value $K$, define a consumption function $\Phi^0(K_t) = X_t$. The required motion on the saddlepath makes $(0, \bar{K}]$ an invariant set for $K_t$, $t \ge 0$.

Next, set $n = 1$. That allows one future TFP jump with hazard $\lambda$. The Euler and budget
equations are

\[ u'(X_t) = \rho (1 - \lambda) (1 + m(A_{t+1})) \cdot u'(X_{t+1}) \]
\[ + \rho \lambda b \left( 1 + m \left( \frac{bA_{t+1}}{Z} \right) \right) \cdot u' \left( Z\Phi^0 \left( \frac{bA_{t+1}}{Z} \right) \right) \]

\[ A_{t+1} = K_t + m(K_t) \cdot K_t + w(K_t) - X_t. \]

Setting \( n = 1 \), Figure 2 presents the phase diagram. Suppose we again have a saddlepath, with movement along the saddlepath characterizing behavior if no shock occurs. If there is a shock, we switch to the previous problem (where \( \Phi^0(.) \) guides expectations). The isocline for the budget equation is as in Fig 1; the isocline for the Euler equation, \( \xi^0(K) \), now has a positive slope. Suppose the saddlepath again has properties (i)-(iii). Define \( \Phi^1(K) = X \) from its height for each \( K \in (0, \hat{K}] \). Now, \( K_{t+1} = A_{t+1} \cdot \omega(z_{t+1}) \). By construction, we again have rational expectations – both initially and after the first shock. And, \( (0, \hat{K}] \) remains an invariant set for \( K_t, t \geq 0 \).

In practice, we derive \( \Phi^n(.) \) from solving a sequence of continuous time two-point boundary value problems. That enables us to take advantage of existing software and to use an algorithm that generates a saddlepath for every \( K \) in the interval \( (0, \hat{K}] \) in a single pass. The pair of equations for the continuous-time problem on iteration \( n \) is

\[ (1 - \gamma) \frac{\dot{X}^n_t}{X^n_t} = m(K^n_t) - \rho_0 - \lambda + \lambda b [X^n_t]^{1-\gamma} \cdot \left[ Z\Phi^{n-1} \left( \frac{bK^n_t}{Z} \right) \right]^{\gamma-1}, \]

\[ \dot{K}^n_t = \eta Y (K^n_t) - \delta K^n_t - X^n_t. \]

We repeat the same procedure on the next iteration, assuming that household beliefs about consumption conditional on the next technology arrival are \( \Phi^n(K) \).

Figure 2 shows the phase diagram for iteration \( n \). The \( \dot{X}^n = 0 \) isocline,

\[ X = \kappa(K) = \eta Y (K) - \delta K, \]

is independent of \( n \). The \( \dot{X}^n = 0 \) isocline for (19) is

\[ X = \xi^{n-1}(K) = Z \cdot \Phi^{n-1} \left( \frac{bK}{Z} \right) \cdot \left[ \frac{\lambda + \rho_0 - m(K)}{\lambda b} \right]^\frac{1}{\gamma}, \]

with \( \xi'(K) > 0 \) and a positive horizontal intercept \( \hat{K} > 0 \) where \( m(\hat{K}) = \lambda + \rho_0 \) and \( \xi(\hat{K}) = 0 \). The trajectory \( K^n_t = K^n [z'] \) moves in cycles within the invariant set \( (0, \hat{K}] \), with each new GPT arrival starting a new cycle. Detrended capital and consumption, \( (K^n_t, X^n_t) \), both rise along the
lower segment of the stable arm while $J_t$ is constant. When a new technology arrives, the detrended capital stock resets to a lower value, $K^n_{t+1} = bA^n_t / Z$, and movement along the saddle path resumes.

We continue iterating in this manner for higher and higher $n$. Suppose the phase diagram continues to exhibit a saddlepath solution with properties (i)-(iii). We anticipate that as $n$ increases and the expected cessation of TFP growth moves farther and farther into the future, the saddlepath $\Phi^n(\cdot)$ will stabilize. We propose checking that outcome by asking whether the sequence of consumption functions $\Phi^n(\cdot)$ is Cauchy. If so, there exists a limit function, and properties (i)-(iii) carry over in the limit. As we show in the next subsection, these are sufficient for existence of a rational expectations equilibrium.

### 3.2 Equilibrium characterization

This section establishes that if the sequence of functions $\Phi^n(\cdot)$ recursively defined above converges uniformly, its limit, $\Phi^*(\cdot)$, can be used to construct a rational expectations equilibrium. We have

**Proposition 3** Suppose that the sequence $\Phi^n$ defined in Section 3.1 is Cauchy and that each saddlepath $\Phi^n : (0, \bar{K}] \rightarrow [0, \infty)$ is (i) non-decreasing, (ii) continuous, and (iii) manifests direct motion towards a unique stationary point $\bar{K}^n$ from every initial condition in $(0, \bar{K}]$. Then there exists a limit function $\Phi^*(K) = \lim_{n \to \infty} \Phi^n(K)$, and an invariant set $\mathbb{K} = [\bar{K}, \bar{K}] \subset (0, \bar{K}]$ such that any trajectory $K[\cdot]$ originating in $\mathbb{K}$ and obeying the aggregate law of motion (10) – with $\Phi(K) = \Phi^*(K)$ – stays in $\mathbb{K}$ for all $t \geq 0$; gross investment is non-negative for all $t \geq 0$; and, the
Figure 3: Equilibrium trajectories for output \( (Y_t) \), capital stock \( (K_t) \), consumption \( (C_t) \) and marginal product of capital \( (M_t) \).

The household wealth trajectory \( k \left[ z^t \right] \) with \( k \left[ z^t \right] = K \left[ z^t \right] \) is consistent with utility maximization in (13) and rational expectations.

**Proof:** See Appendix.

**Equilibrium Trajectories** Proposition 3 and (3) imply that the aggregate capital stock \( K_t \) monotonically converges to a steady state \( Z^J \bar{K} \) while \( J \) is constant. The equilibrium trajectory for \( K_t \) then has a “saw-tooth” pattern – shown on the upper left panel of Figure 3 – with intervals where \( K_t \) is rising punctuated by abrupt drops in the market value of the capital stock by factor \( b \). The marginal product of a dollar of capital, \( M_t \), rises when \( K_t \) falls.

The interval \( [t_J, t_{J+1}] \) represents the diffusion phase for technology \( J \): the share of \( K_{J,t} \) in the total capital stock rises as new investment builds up \( K_{J,t} \), and older vintages are left to decay from wear and tear. Convergence to the steady state is interrupted at time \( t_{J+1} \) when technology \( J + 1 \) arrives. After an initial drop, the value of capital stock starts another upward climb, this time towards a higher steady state, \( Z^{J+1} \bar{K} \).

The equilibrium trajectory for \( C_t \) follows a similar pattern – consumption is rising while \( J \) stays constant, but it abruptly falls when \( J \) increases. Households cut consumption in response to new investment opportunities signaled by a rise \( M_t \) upon a new GPT arrival. In our formulation so far,
$\gamma_{t+1}$ does not change at all when $z_{t+1} = 1$, because the newest vintage has not yet been built, and production as before remains possible. In aggregate production function (3), a new GPT arrival causes a decrease in the value of $K_{t+1}$ that is fully offset by the increased productivity of the new GPT. However, in our calibrations below, we consider a slight generalization of the model in which output may fall as a new GPT arrives.

### 3.3 Risky and risk-free rates

Consider the risk-free rate. For expositional convenience, set the period length to $h$ and present expressions with $h \to 0$. Let $\rho_0 = -\ln \rho$ be the instantaneous discount rate corresponding to the time discount factor $\rho$. The riskless rate between dates $t$ and $t + h$, $r_{t,t+h}$ obeys the Euler equation

$$u'(C_t) = (1 + (r_{t,t+h} - \rho_0)h) \cdot \mathbb{E}[u'(C_{t+h})].$$

Use the notation $\dot{C}_t$ for the rate of change in consumption when there is no GPT arrival at date $t$, and let $\Delta u'_{t+h} = u'(C_{t+h}) - u'(C_t)$ denote the (upward) jump in marginal utility when a new GPT arrives at date $t+h$. Then

$$\mathbb{E}[u'(C_{t+h})] \approx (1 - \lambda h) \left( u'(C_t) + u''(C_t) \dot{C}_t \cdot h \right) + \lambda h \cdot u'(C_{t+h}) \approx u'(C_t) + u''(C_t) \dot{C}_t \cdot h + \lambda h \cdot \Delta u'_{t+h},$$

and the first-order approximation for the Euler equation is

$$u'(C_t) \approx (r_t - \rho_0)h \cdot u'(C_t) + u'(C_t) + u''(C_t) \dot{C}_t \cdot h + \lambda h \cdot \Delta u'_{t+h}.$$ 

Collecting the terms and taking the limit $h \to 0$, the risk-free rate is

$$r_t = \rho_0 + (1 - \gamma) \frac{\dot{C}_t}{C_t} - \lambda \frac{\Delta u'_t}{u'(C_t)}. \quad (21)$$

The average return on the risky asset can be derived in a similar fashion. The Euler equation on the interval from $t$ to $t+h$ is

$$u'(C_t) = (1 - \rho_0 h) \mathbb{E}\left[(1 + R_{t,t+h}) u'(C_{t+h})\right]. \quad (22)$$

The risky return over the time interval $[t, t+h]$ is

$$1 + R_{t,t+h}h = \begin{cases} 
1 + m_t h, & \text{with prob. } 1 - \lambda h, \\
b + m_t h & \text{with prob. } \lambda h.
\end{cases} \quad (23)$$

---

9The analog of (11) is $\rho_0 > \lambda (Z^\gamma - 1)$. 

16
We have
\[ \bar{R}_t = \mathbb{E} R_{t,t} = m_t - \lambda (1 - b) \] (24)
for the average risky rate of return on capital. Repeating the steps above for (22)-(24) yields
\[ \bar{R}_t = \rho_0 + (1 - \gamma) \frac{\dot{C}_t}{C_t} - \lambda b \frac{\Delta u'_t}{u'(C_t)} \] (25)

The last terms in (21) and (25) are non-standard: they register the (large) effects of new GPT arrivals on marginal utility. As will become clear when we present quantitative results, these terms contribute to both a sizeable equity premium and a low risk-free rate in our model. If \( \Phi^* \) is the equilibrium aggregate consumption function, note that in the limit \( h \to 0 \) we have
\[ \frac{\Delta u'_t}{u'(C_t)} = \frac{u'(Z \Phi^* (\frac{bK_t}{Z}))}{u'(\Phi^* (K_t))} - 1. \]

The expected risk premium is
\[ \bar{R}_t - r_t = \lambda (1 - b) \frac{\Delta u'_t}{u'(C_t)}. \]

It equals the average rate of obsolescence, \( \lambda (1 - b) \), times the change in marginal utility when a GPT arrives. As explained above, the risk premium arises from the negative co-movement of \( R_t \) and the marginal utility of consumption. An innovation in the frontier TFP lowers the cost of capital services, inflicting a capital loss on equity holders. At the same time, consumption drops as households increase their saving to take advantage of the new investment opportunities. Marginal utility then rises just as returns on past equity holdings are low, fueling the equity premium.

4 Calibration

We have seven parameters to calibrate. There are two for the TFP process, \( \lambda \) and \( \theta; \) two for income shares, \( \alpha \) and \( \beta; \) two for preferences, \( \gamma \) and \( \rho; \) and, one for the rate of wear and tear depreciation, \( \delta \). Our calibration process depends mainly on national accounts data. In particular, it does not target financial market moments that we use for quantitative assessments of the model.

TFP process To characterize the evolution of the TFP frontier, consider the growth factor for output per worker associated with a new GPT, \( Z = \theta^{1/(1-\alpha)} \), and the frequency of GPT arrivals, \( \lambda \). Neither is directly observable. However, Proposition 1 shows that equity prices drop at the arrival
of a new GPT and the magnitude of the decline provides information about the magnitude of $Z$:

$$b = Z^{-\frac{1-\alpha}{\alpha}}. \tag{26}$$

In in model of Section 2, the arrival of a new GPT does not reduce output, and production with older vintages never ceases. In practice, disruptions accompanying the arrival of a new GPT may cause output to fall. For example, there may be frictions in transforming the investment-good sector to produce the new vintage of capital, financial repercussions from the stock market crash, and/or problems in reaching new market-clearing prices (e.g. Basu et al., [2006]). We can adjust our framework to accommodate such factors by assuming that only a fraction $\Delta$ of the beginning capital stock remains operational. In our simulations, for example, the disruption from a new GPT leads to an output loss of 5 percent.\(^{10}\) Set $\Delta = (0.95)^{1/\alpha}$. The combined effect of GPT arrival is then to reduce the price of capital services by a factor $b$ and to reduce the physical quantity of capital by a factor $\Delta$, with the resulting stock market revaluation factor $\hat{b} = \Delta \cdot b$. We use the estimate of the revaluation factor $\hat{b} = 0.3866$ from Laitner and Stolyarov [2003], and calibrate $Z$ from

$$\hat{b} = \Delta \cdot Z^{-\frac{1-\alpha}{\alpha}}. \tag{27}$$

If we further require consistency with the observed long-run average productivity growth rate, say, $g = 0.021$, the model relates $Z$, $g$, and $\lambda$ as follows. Let $S$ denote a long time horizon. The empirical growth factor for per capita output over that horizon is $e^{g \cdot S}$. The model’s forecast of the average factor is $E_J[Z^J]$.\(^{11}\) Equating the two gives a calibration equation that relates $\lambda$ and $Z$:

$$e^{g \cdot S} = E_J[Z^J] = e^{\lambda S(Z-1)} \iff Z = \frac{g + \lambda}{\lambda}. \tag{28}$$

We present two alternative calibrations. Calibration 1 sets $\alpha = 0.3$ and solves (27)-(28) for $Z$ and $\lambda$. Calibration 2 sets $\lambda = 0.04$ from historical data and solves (27)-(28) for $Z$ and $\alpha$.\(^{12}\)

**Depreciation Rate** In the model, the depreciation rate consists of two components: physical wear and tear, $\delta$, and obsolescence, $\bar{\delta}$. The latter registers the long-run effect of abrupt depreciation episodes accompanying new GPT arrivals. The model relates the obsolescence component to $\lambda$ and

---

\(^{10}\)By way of comparison, a calibration in Gabaix (2011) assumes a 20 percent permanent output loss from a disruptive event.

\(^{11}\)With $h \to 0$, GPT arrivals have a Poisson density. Hence,

$$e^{g \cdot S} = \sum_{j \geq 0} \frac{(\lambda \cdot S)^j \cdot e^{-\lambda \cdot S}}{j!} \cdot Z^j \iff e^{(g+\lambda) \cdot S} = \sum_{j \geq 0} \frac{(\lambda \cdot Z \cdot S)^j}{j!}.$$  

Since the last term is the power series for $e^{\lambda Z \cdot S}$, we have $e^{g \cdot S} = e^{\lambda S(Z-1)}$.

\(^{12}\)Gordon [2012] identifies 5 GPTs over 125 years, or, one new GPT every 25 years, on average. With our assumption of Poisson arrival, this implies $1/\lambda = 25$.  

18
\( \hat{b} \) as follows:

\[
e^{-\delta s} = \mathbb{E}_t \left[ \hat{b}^t \right] = e^{\lambda s (\hat{b} - 1)} \iff \delta = \lambda \left( 1 - \hat{b} \right).
\]

If \( d = \delta + \bar{\delta} \) is the total depreciation rate, then

\[
\delta = d - \lambda \left( 1 - \hat{b} \right).
\]

As explained above, the model excludes residential assets and housing services; therefore, we determine \( \delta \) from this formula and the NIPA time-series average depreciation rate on non-residential capital, \( d = 0.0752 \).

**Leverage** In our model, the Modigliani-Miller theorem applies; consequently, the mix of debt and equity financing does not affect the theoretical analysis. The mapping from the theoretical risk premium to the empirical equity premium does, nevertheless, depend on the debt-equity ratio because equities claim only a portion of capital income in practice (e.g., Barro (2006), p.843). Letting \( \nu \) be the share of debt financing, the model’s overall return on capital is a weighted average of the rate of return on equities, \( R^S_t \), and debt:

\[
R_t = \nu r_t + (1 - \nu) R^S_t.
\]  

(29)

Accordingly, the return on levered equity is

\[
R^S_t = \frac{R_t - \nu r_t}{1 - \nu},
\]

and the leverage-adjusted equity premium is

\[
E_t = R^S_t - r_t = \frac{R_t - r_t}{1 - \nu}.
\]

Using data from the U.S. Flow of Funds on debt-equity ratios for non-financial corporations, we set \( \nu = 0.40 \).\(^{13}\)

**Type-I Labor Income Share** The distribution of labor income between type-I and type-II workers influences the relationship between type-I consumption, \( C_t \), and aggregate consumption, \( \bar{C}_t \):

\[
C_t = C_t + (1 - \beta) (1 - \alpha) \bar{y}_t.
\]

Type-I agents receive all of the economy’s capital income and fraction \( \beta \) of labor income. A GPT arrival has little or no impact on the latter. Hence, a higher \( \beta \) tends to smooth the overall income

\[^{13}\text{We calculate } \nu_t \text{ from the debt-to-equity ratio for non-financial corporations, US. Flow of Funds series ID FL104104016. See also Boldrin et al. [1995] and Papanikolaou [2011].} \]
flow of type-I households. We calibrate $\beta$ from the income share of the top 5% of individuals, a group that we associate with type-I agents. In the past 100 years, income concentration followed a U-shape pattern, with the top 5% income share moving within a 22-40% range (Saez (2017), Figure 2). Based on this, we pick the top 5% income share of $1/3$, corresponding to 2007 data from Kennickell (2009). If $Y$ and $D$ are non-residential GDP and depreciation, respectively, the calibration equation for $\beta$ is

$$\frac{1}{3}[Y - D] = [\alpha \cdot Y - D] + \beta \cdot (1 - \alpha) \cdot Y.$$  

Using $\alpha$ set in the calibration and 2007 data for $Y$ and $D$, we have $\beta = 0.1563$ for calibration 1 and $\beta = 0.1269$ for calibration 2.

**Preference parameters** Our choice of $\gamma$ is disciplined by consistency with aggregate consumption data. A lower IES (higher CRRA) moderates movement in consumption when technology shocks occur. For the range of $\gamma$ that we are considering, a lower IES also makes the equity premium smaller. In each calibration, we choose a $\gamma$ at the higher end of the IES range where model-generated fluctuations in aggregate consumption per capita are in line with those observed during the 128 year period 1889-2017. This procedure selects $\gamma = -0.5$ (IES of 2/3) for calibration 1, and $\gamma = -0.25$ (IES of 0.8) for calibration 2. Below we report the corresponding frequencies of large aggregate consumption declines. The rate of time preference, $\rho_0$, is set from the value the long-run “Ramsey” interest rate, $\bar{r} = \rho_0 + (1 - \gamma) g$. The standard range for $\bar{r}$ in the macroeconomics literature is 4-5 percent. Accordingly, we set $\rho_0 = 0.045 - (1 - \gamma) g$ in the calibrations.

**Simulations** Each new GPT determines the frontier technology for the ensuing interval. Using a random number generator, we draw 1000 sample interval durations from the exponential distribution. Starting from the initial condition $K = \bar{K}$, we run the model through the set of intervals once to set an initial $K_0$. Starting from the latter, we run through the set of intervals again, calculating trajectories for comparisons with the data. The continuous-time simulations are averaged over 1-year segments to generate discrete data points at an annual frequency. We calculate means and standard deviations using the latter.

**Data description** We evaluate the model by comparing moments from the model’s simulated trajectories with their data analogs. We use annual data covering the period 1871-2017. The stock price and earnings data are from the S&P composite index data in Shiller (1989, Ch 26, Series 1, 3), updated to 2017. The series for real per-capita consumption is taken from Shiller (1989, Table 26.2 series 9) for 1889-1929 and spliced with NIPA after 1929. All the rates calculated from the data are geometric rates, $\ln(1 + R)$, which seem the most consistent with instantaneous rates

---

14 Using 2007 data, measured GDP was 14,477.6 bn/yr, from which we subtract services of residential housing (1480.0 bn/yr) and investment in residential housing (688.7 bn/yr), leaving $Y = 12,308.9$ bn/yr. NIPA depreciation was 1865.0 bil/yr, from which we subtract depreciation on residential housing, 406.7 bn/yr.
in our continuous-time approximation solution method. (Geometric rates are lower than percent changes, and the difference $R - \ln (1 + R)$ is rising with $|R|$.) We define the risk-free rate as the real annualized rate of return on prime 4-6 month commercial paper as reported in Shiller (1989, Ch 26, Series 4).\footnote{The Federal Reserve Board discontinued its 6-month commercial paper rate series in August 1997. After this date, Shiller uses the 6-month certificate of deposit rate, secondary market, FRED series CD6NRRNJ, with FDIC as a data source.} For the 1947-2017 sub-sample, we use additional data on capital structure to construct moments for the “unlevered” return on equity, the direct analog of our $R_t$, using the empirical time series for $R_t^S$, $r_t$ and $\nu_t$.

## 5 Model assessment

Table 1 reports on simulations of the model. Overall, we find a reassuring degree of agreement between the model and the data. In particular, the means and coefficients of variation for the equity premium, as well as the Sharpe ratio, are consistent with the data.

[Table 1 here]

Another empirical success of the model is to replicate both the level and the volatility of the risk-free rate. Section 5.2 discusses this in detail. In general, the alignment between the model and the data is somewhat better over the longer sample.

Comparisons of earnings between the model and the data are not straightforward. Earnings in the model are the difference between the marginal product of one dollar’s worth of capital and the wear and tear rate of depreciation. In the data, they measure the marginal product less total depreciation. Moreover, the model evaluates depreciation in terms of the market value of capital, whereas the data tends to use the book value. We calibrate that a fraction $\nu$ of each dollar’s worth of capital is debt financed. The data will subtract $\nu$ times the rate of return on corporate debt, as it reports earnings yield per dollar of equity. Our calculation of $m_t$ for the model does not do so. If we subtract interest on debt from earnings using $\nu = 0.4$ and the model’s riskless rate, the alternative measure of earnings yield, $m_t - \nu r_t$, would be 7.97 percent for calibration 1 and 6.74 percent for calibration 2. According to the model, the riskless rate rises in the aftermath of the arrival of a new GPT. Thus, the simulated standard deviation of $m_t - \nu r_t$ will be less than $\sigma (m_t)$ reported in Table 1.

### 5.1 Equity premium

As noted, the equity premium for calibration 1 (calibration 2) is 5.30%/yr (4.45%/yr) in Table 1. A surprising result is that a sizeable equity premium obtains for low coefficients of relative risk.
aversion (CRRA). Note that the model’s risk premium can be non-monotone in the risk aversion parameter $1 - \gamma$. Consider a linear approximation for the jump in marginal utility after a GPT shock

$$\frac{\Delta u'_t}{u'(C_t)} \approx -(1 - \gamma) \frac{\Delta C_t}{C_t}.$$ 

A higher CRRA, $1 - \gamma$, has opposing effects on the marginal utility change term. On the one hand, risk aversion amplifies the influence of consumption growth rate on marginal utility. On the other hand, higher risk aversion (i.e., a lower IES) makes households less willing to increase saving in response to new investment opportunities brought by a GPT arrival. Consequently, $\frac{\Delta C_t}{C_t}$ falls as CRRA rises. The latter, IES effect is especially influential in the model. In simulations for $\gamma \in [-1.0, 0.5]$, the equity premium first rises, then falls. The highest premium for calibration 1 (5.57%/yr) occurs at $\gamma = 0.00$, and for calibration 2 (4.71%/yr) at $\gamma = 0.25$. We chose slightly smaller values of $\gamma$ for Table 1 to match statistics on aggregate consumption volatility — see below.

The magnitude of the equity premium positively depends on the size of TFP innovations, as the extent of capital losses and the drop in consumption upon a new GPT’s arrival both depend on $Z$. Our solution algorithm preserves the influence of the model’s nonlinearities, so that, for instance, convex marginal utilities can amplify the effect of large shocks. We expect less frequent but larger shocks to yield a higher equity premium. That is borne out: if we halve (double) $\lambda$ in calibration 1 – computing $Z$ from (28), preserving $g = 0.021$, and calculating a new $b$ in (27) with $\alpha = 0.30$ – the equity premium is 6.40%/yr (4.47%/yr).

A contemporaneous decline in output upon GPT arrival can also have a large effect. Our calibrations assume a 5% drop in output that we add to the baseline model by setting $\Delta < 1$ in (27). As explained, we think of $\Delta$ as reflecting frictions in adjusting to the newly arrived technology, financial repercussions from the stock market crash, and a decentralized economy’s problems in reaching new market-clearing prices. The output decline reduces type-I agents’ resources at a time when their marginal utility is high, compounding their discomfort from portfolio losses. Importantly, the decline reduces both their labor and capital incomes. That tends to increase the equilibrium equity premium. If we halve (double) the output decline accompanying a new GPT in calibration 1 – adjusting $\alpha$ in (27) but maintaining other parameters unchanged – the equity premium is 4.86%/yr (6.24%/yr). The model’s simplicity suggests that it can accommodate multiple shocks, and future work might, for instance, consider the possible effects of preference shifts or oil-price changes.

In contrast to many papers in the literature, our type-I agents have a degree income insurance from their labor earnings. That tends to reduce the model’s equilibrium equity premium because earnings do not suffer steep reductions when a new GPT arrives. Earnings then can buffer the impact of capital losses on equities. The effect can be surprisingly large. If we halve (double) $\beta$ in

---

16 The calibration for $\alpha$ changes to 0.2796 (0.3420).
calibration 1, the simulated equity premium is 6.19%/yr (4.21%/yr).

Comparisons with consumption data In their response to Rietz (1988), Mehra and Prescott (1988) suggested disciplining predictions of “rare disaster” models with consumption data. As a validity check, we compare our model’s predictions on consumption volatility with per-capita data on consumption 1889-2017.\(^{17}\)

A GPT arrival raises the current and future marginal products of capital, inducing type-I households to cut back on consumption to take advantage of the new investment opportunities. Resulting fluctuations in aggregate consumption can be very large if the IES is low. The calibrations in Table 1 limit \(\gamma\) to match the observed frequency of declines in aggregate consumption exceeding 5 (10) percent, as reported in Table 1.\(^{18}\)

We can separately look at consumption volatility for type-I agents. Their per capita consumption, \(C_t\), moves more than aggregate per capita consumption, \(\bar{C}_t\), because they bear all of the economy’s asset-return risk. For type-I agents, the model predicts a maximum (i.e., once in 20,000 years) consumption drop of about 51% for calibration 1, and about 61% for calibration 2. Though these magnitudes are large, changes in stock market wealth may affect type-I agent expenditures on discretionary luxury goods rather than basic consumption. Ait-Sahalia et al. (2004) document luxury-good consumption over a 40-year period and show that expenditures in that category are an order of magnitude more volatile than aggregate NIPA personal consumption expenditure – and that their covariance with stock returns is positive. For example, Ait-Sahalia et al. (2004, Fig 3) reports that sales of luxury automobiles and charitable contributions both experienced one-year drops of over 60 percent during the 1961-2001 period, and that other categories of luxury goods show multiple episodes of 20-30 percent year-on-year sales declines.\(^{19}\)

5.2 Risk-free rate

The simulations of Table 1 display relatively low risk-free interest rates. To explain their magnitudes, we decompose the expression for the average risk-free rate into a standard term, \(\rho_0 + (1 - \gamma) g\), and a (negative) residual term that depends on large changes in marginal utility that accompany new GPT arrivals. The risk-free rate is determined in equation (21),

\[ r_t = \rho_0 + (1 - \gamma) \frac{\dot{C}_t}{C_t} - \lambda \frac{\Delta u_t'}{u'(C_t)}, \]

---

\(^{17}\)We use consumption series from Shiller (1989, Table 26.2 series 9) for 1889-1929, and NIPA data after 1929.

\(^{18}\)Another way to compare consumption fluctuations between the model and data is as follows. We partition our 20,000 years simulation into 100-year intervals and calculate the peak consumption drop within each interval. The data 1889-2017 shows a maximum yearly consumption per capita decline of 10.03 percent. The model-generated peak consumption drops are less than 10.03 percent for about 55 % of the 100-year intervals in calibration 1 and for about 36 % of the 100-year intervals in calibration 2.

\(^{19}\)See also Mankiw and Zeldes [1991], Vissing-Jorgensen [2002], Attanasio et al. [2002], and others.
Figure 4: Change in marginal utility after a large drop in consumption

where \( \Delta u_t' = u'(C_{t+h}) - u'(C_t) \) is the change in marginal utility when a GPT arrives at date \( t+h \). As in Section 3.3, \( \Delta C_t = C_{t+h} - C_t \). The finite change \( \Delta u_t' \) can be written as the sum of its linear approximation and higher-order terms, \( \tilde{u}_t' \):

\[
\Delta u_t' = (\gamma - 1) u'(C_t) \frac{\Delta C_t}{C_t} + \tilde{u}_t'.
\]

Since \( u'(\cdot) \) is a convex function, \( \tilde{u}_t' \) is positive, as shown in Figure 4. Using the above expression for \( \Delta u_t' \) in (21) yields

\[
r_t = \rho_0 + (1 - \gamma) \frac{\dot{C}_t + \lambda \Delta C_t}{C_t} - \lambda \frac{\tilde{u}_t'}{u'(C_t)} \tag{30}
\]

Our simulations take the first right-hand side term, \( \rho_0 \), to be a positive parameter. Over a long horizon \( S \), the second term encompasses smooth growth between GPT arrivals, at velocity \( \dot{C}_t \), and \( \lambda S \) new GPT arrivals, each causing a consumption decline \( \Delta C_t \). Since the model displays balanced growth in the long run, the average value of the second term is \((1 - \gamma)g\). The third term is negative.

Eq (30) shows that our model can generate a low riskless rate for two reasons. First, the model’s structure allows a sizeable equity premium with a CRRA of less than 1.5. Therefore, the \((1 - \gamma)g\) term can be relatively small. In fact, \((1 - \gamma)g\) is less than 300 basis points in our calibrations. Second, the model’s large shocks tend to make the last term large in absolute size (see Figure 4). In calibration 1, it contributes \(-192\) basis points to the average risk-free rate; in calibration 2, \(-180\) basis points.
5.3 Return volatility

Empirically, stock price movements do not seem to be well-explained by subsequent dividend changes (e.g., Shiller [1981]). This is a puzzle (i.e., the “volatility puzzle”) if the shocks that move stock prices are assumed to come from news about future dividends. Our model suggests an interpretation, however.

The model has a general equilibrium framework in which a GPT shock jointly affects stock prices, dividends, and stochastic discount factors. The arrival of a new technology lowers the supply price of capital, causing stock prices to drop from $1$ to $b$, as specified in Proposition 1. Yet, the marginal product per dollar’s worth of $K$ subsequently rises — recall Figure 3. If we expect a future drop in dividends to justify the sharp fall in stock prices at a GPT’s arrival, we will be disappointed. On the contrary, the model predicts that a punctuated technology shock will cause stock prices and returns (per dollar) on capital to move in opposite directions. There is no contradiction of rationality; the model’s definition of equilibrium requires the stock price to equal the present value of dividends. The intertemporal pricing kernel absorbs the disparate movements in stock prices and capital’s earnings through an adjustment to current consumption — recall Section 5.1. As it turns out, the mechanism allowing stock prices and dividends to move in opposite directions is, therefore, the same as the one driving the equity premium.

[Table 2 here]

Table 2 presents long-term U.S. data on correlations between time $s = 0$ stock price changes and past and future dividends and earnings. The model predicts a negative correlation with earnings for $s > 0$, but no relation for $s < 0$. That seems consistent with the data. As we do not have a separate model of dividends, focus on columns 3-4. We see that the correlation of current capital gains with lagged dividend or earnings yields is low, especially in the longer time series. For $s \geq 1$, the correlation reverses the sign, and it stays negative over a 10-year horizon.

Table 2 shows a positive correlation between stock prices and contemporaneous earnings. Our benchmark specification of Sections 2-3 would predict a zero correlation. However, Section 4 calibrates a slightly abridged version of the model that assumes a new GPT renders some existing capital useless – compare (26) and (27). Then earnings would fall at the same moment as stock prices, explaining a time $s = 0$ positive correlation in the table.

5.4 Return predictability

The classic (Fama [1965]) “random walk” hypothesis holds that efficient financial markets should incorporate all available information into the current stock price, making expected stock returns
constant over time. In the data, however, stock returns can be forecasted with current dividend yields (see, e.g. Campbell and Shiller [1988], Cochrane [2008, Table 1] and Gabaix [2012, Table IV], among many others). Predictable stock returns are considered puzzling because it is not clear why market participants are not trading on available information. Our model’s explanation for the puzzle is based on the gradual nature of technology diffusion. Agents’ desire to smooth consumption prevents them from taking full advantage of the higher stock returns that follow a new GPT arrival. As a result, high returns persist for some time after a GPT shock, and they are predictable even at moderately long horizons.

In the model, expected return per dollar invested in equities depends upon $K_t$:

$$\bar{R}_t = m(K_t) - \lambda \cdot (1 - b).$$

In the time series, the model produces a sequence of episodes with gradual stock run-ups followed by steep declines. Within each bull market episode, the price-earnings ratio,

$$\frac{P_{jt}}{MPK_{jt} - \delta \cdot P_{jt}} = \frac{1}{m_t},$$

gradually rises, then falls abruptly when a new technology arrives. A regression of a forward return per dollar invested on, say, the logarithm of the current dividend-price ratio would yield a positive coefficient (as in Cochrane and Gabaix [op. cit.]). In a sequence of price-earnings ratios, the model predicts a pattern of positive autocorrelations, falling with the time horizon. The data in Table 2, columns 3-4, show such a pattern. Over the 1871-2017 time period, the autocorrelations of the price-earnings ratios at 1, 3, 5, and 10-year horizons are 0.59, 0.38, 0.31, and 0.26, respectively. The same statistics for the model are higher, because the model’s focus on low-frequency shocks tends to overstate persistence. Using an annualized time series from calibration 1, for instance, we find autocorrelations of the price-earnings ratio of 0.93, 0.81, 0.70, and 0.47 at 1, 3, 5, and 10-year horizons.

The model’s sequence of bull and bear markets also seems broadly consistent with evidence. Albuquerque et al. [2015, Table 1] reports that in U.S. data, bull markets have an average duration of 14.8 years, and bear markets are much shorter, with 3.2 years average duration. In our model, of course, bear markets are instantaneous and bull markets have average duration $1/\lambda$ (18.89 years in calibration 1, and 25 years in calibration 2). Albuquerque et al. [2015, Table 4] finds that stock returns are strongly positively correlated with output and consumption growth within each episode. That, too, is in agreement with our model: as the stock market is rising (when growth proceeds up the model’s saddlepath toward the stationary point), stock returns, consumption, and output all depend upon $K_t$, which is rising.
6 Conclusion

This paper studies disruptive technological change as a source of macroeconomic risk. Our model features large, but infrequent technology shocks. Each sharply improves society’s frontier technology, and causes sweeping change in the economy. We associate the shocks with arrivals of new general purpose technologies. Two attributes of GPT shocks are key to our analysis. First, a new type of capital has to be built to realize the productivity advantage of a new technology. Second, each shock reduces the value of existing assets (due to obsolescence), though, at the same time, raises future returns on new investment.

We focus on the model’s implications for asset pricing. We show that GPT shocks give rise to economic mechanisms that can potentially explain four long-standing asset pricing puzzles: a sizable equity premium, a low risk-free rate, and stock returns that are both volatile and predictable. Our model provides a new perspective on the puzzles and show that GPT shocks can provide a unifying explanation for all four.

The equity premium arises from the negative co-movement of stock returns and the marginal utility of wealth when a technology shock occurs. A new GPT lowers the production cost for capital services, and the price of existing assets falls in step, depressing stock returns. At the same time, households respond to new investment opportunities brought by the new technology by cutting their consumption to increase their saving. Since the marginal utility of wealth then rises when stock returns are low, households demand a positive risk premium. Importantly, the mechanism does not rely on fluctuations of output to generate a correlation between consumption and stock returns.

The risk-free rate level is low for a related reason. Households anticipate a future possibility of a new technological change, which raises their expected future marginal utility of wealth, lowering the risk-free rate.

The paper suggests that a mechanism driving stock return volatility is related to the simultaneous, differential impact of GPT shocks on current and future returns. The conventional view holds that stock prices respond to news about dividends. A GPT shock, however, affects both investment good supply prices and future dividends, and it moves them in opposite directions. If stock prices crash at a time when future dividend yields rise, discount rates have to rise sharply. Nevertheless, the same mechanism that generates the equity premium – namely, the negative correlation between capital gains and stochastic discount rates – allows the necessary discount rate change to be an equilibrium.

Lastly, our explanation for return predictability is based on gradual diffusion of new GPTs. The conventional view holds that if financial markets are efficient, agents should be able to trade on any new information, which should leave stock returns unpredictable. Our explanation for the puzzle
is that agents’ desire to smooth consumption prevents them from taking immediate advantage of the higher stock returns associated with a new technology. Instead, the gradual nature of diffusion leaves stock returns high, and predictably so, over moderately long periods.

Our framework builds on a standard, workhorse macroeconomic model. Production is neoclassical, with constant returns to scale in labor and capital. Consumer utility functions are isoelastic, and the coefficient of risk aversion in the calibrated examples is between 1 and 1.5. The model’s overall parametrization is parsimonious, and its calibration does not target financial market variables. The model features a single source of aggregate risk, and this is, of course, a simplification. The model’s tractability suggests that generalizations to several types of shocks may be feasible, potentially adding both richness and realism to the basic framework.
References


<table>
<thead>
<tr>
<th></th>
<th>Model Calibration 1&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Model Calibration 2&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Data 1871-2017</th>
<th>Data 1947-2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock return, $\bar{R}^S$</td>
<td>7.88</td>
<td>7.15</td>
<td>6.55</td>
<td>7.12</td>
</tr>
<tr>
<td>Risk-free rate, $\bar{r}$</td>
<td>2.58</td>
<td>2.70</td>
<td>2.35</td>
<td>1.25</td>
</tr>
<tr>
<td>Equity premium, $E = \bar{R}^S - \bar{r}$</td>
<td>5.30</td>
<td>4.45</td>
<td>4.20</td>
<td>5.87</td>
</tr>
<tr>
<td>Sharpe ratio, $\bar{E}/\sigma(E)$</td>
<td>0.22</td>
<td>0.22</td>
<td>0.24</td>
<td>0.38</td>
</tr>
<tr>
<td>Rate of return on equity, unlevered, $\bar{R}$</td>
<td>5.75</td>
<td>5.37</td>
<td>-</td>
<td>5.70</td>
</tr>
<tr>
<td>Earnings yield, $\bar{m}$</td>
<td>9.00</td>
<td>7.82</td>
<td>7.32</td>
<td>6.92</td>
</tr>
</tbody>
</table>

**Aggregate consumption declines**

| Probability of a drop of over 5 percent | 2.89 | 2.54 | 2.34 | 0.00 |
| Probability of a drop of over 10 percent | 0.55 | 0.83 | 0.78 | 0.00 |

**Standard deviations**

| Stock return, $\sigma(\bar{R}^S)$ | 22.51 | 19.53 | 16.99 | 15.63 |
| Risk-free rate, $\sigma(\bar{r})$ | 7.33  | 6.50  | 6.26  | 2.93  |
| Equity premium, $\sigma(E)$       | 23.66 | 20.47 | 17.32 | 15.56 |
| Rate of return on equity, unlevered, $\sigma(\bar{R})$ | 13.82 | 12.05 | -     | 10.38 |
| Earnings yield, $\sigma(\bar{m})$ | 7.19  | 6.42  | 2.73  | 2.98  |

**Coefficients of variation, $\sigma(x)/\bar{x}$**

| Stock return                   | 2.86  | 2.73  | 2.59  | 2.19  |
| Risk-free rate                 | 2.84  | 2.41  | 2.66  | 2.34  |
| Equity premium                 | 4.47  | 4.60  | 4.12  | 2.65  |
| Rate of return on equity, unlevered | 2.40  | 2.24  | -     | 1.82  |
| Earnings yield                 | 0.80  | 0.82  | 0.37  | 0.43  |

Table 1. Model-data comparisons. Units for rates and probabilities are percent.

a. Calibration 1 parameter values: $\alpha = 0.3, \beta = 0.1563, \hat{b} = 0.3866, \Delta= 0.8428, Z = 1.3966, \delta = 0.0427, \delta = 0.0325, T = 18.885, \lambda = 0.053, \nu = 0.4$.

b. Calibration 2 parameter values: $\alpha = 0.3449, \beta = 0.1269, \hat{b} = 0.3866, \Delta= 0.8618, Z = 1.5251, \delta = 0.0507, \delta = 0.0245, T = 25, \lambda = 0.04, \nu = 0.4$. 
<table>
<thead>
<tr>
<th>( s )</th>
<th>( \text{Corr}(\ln(P_{t+1}/P_t), d_{t+s}/P_{t+s}) )</th>
<th>( \text{Corr}(\ln(P_{t+1}/P_t), e_{t+s}/P_{t+s}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>0.03</td>
<td>0.10</td>
</tr>
<tr>
<td>-1</td>
<td>0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>0</td>
<td>0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>1</td>
<td>-0.24</td>
<td>-0.11</td>
</tr>
<tr>
<td>2</td>
<td>-0.17</td>
<td>-0.09</td>
</tr>
<tr>
<td>3</td>
<td>-0.09</td>
<td>-0.07</td>
</tr>
<tr>
<td>4</td>
<td>-0.14</td>
<td>-0.13</td>
</tr>
<tr>
<td>5</td>
<td>-0.10</td>
<td>-0.20</td>
</tr>
<tr>
<td>10</td>
<td>-0.18</td>
<td>-0.22</td>
</tr>
</tbody>
</table>

Table 2. Correlation of capital gains with past and future dividends and earnings yields
Appendix

Proof of Proposition 1: Let $L_{jt}^*, l_{jt}^* \cdot \bar{t}_{jt}$ denote the solution to (1) and let

$$Y_{jt} = \theta_0 \cdot [\theta]^j [K_{jt}]^a [L_{jt}]^{1-a}.$$  

(31)

**Step 1** Show that $l_{jt}^* = \bar{t}_{jt}, \sum_{j \leq h} L_{jt}^* = 1$ and that $Y_{jt}/L_{jt}^* = y_t$ is independent of $j$. Output maximization in (1) requires that the marginal product of labor of each type be equated across vintages:

$$\beta (1 - \alpha) \frac{Y_{jt} L_{jt}^*}{L_{jt}^* l_{jt}^*} = W_t,$$

(32)

$$(1 - \beta) (1 - \alpha) \frac{Y_{jt} L_{jt}^*}{L_{jt}^* l_{jt}^*} = \bar{W}_t.$$  

(33)

Summing (32) and (33) over $j$ relates $W_t$ and $\bar{W}_t$ to aggregate output:

$$\beta (1 - \alpha) Y_t = \beta (1 - \alpha) \sum_{j \leq h} Y_{jt} = W_t \sum_{j \leq h} l_{jt}^* = W_t$$

(34)

$$(1 - \beta) (1 - \alpha) Y_t = (1 - \beta) (1 - \alpha) \sum_{j \leq h} Y_{jt} = W_t \sum_{j \leq h} \bar{t}_{jt} = \bar{W}_t.$$  

It follows that

$$\frac{W_t}{W_t} = \frac{\beta}{1 - \beta}.$$  

Dividing (32) by (33) gives

$$\frac{W_t \cdot l_{jt}^*}{W_t \cdot \bar{t}_{jt}^*} = \frac{\beta}{1 - \beta} \Leftrightarrow l_{jt}^* = \bar{t}_{jt}^*.$$  

It follows that $L_{jt}^* = [l_{jt}^*]^\beta \cdot \bar{t}_{jt}^{1-\beta} = l_{jt}^*$ and $\sum_{j \leq h} L_{jt}^* = \sum_{j \leq h} l_{jt}^* = 1$. Since $L_{jt}^* = l_{jt}^*$, (32) implies that

$$Y_{jt}/L_{jt}^* = y_t.$$  

(35)

**Step 2** Show that

$$\theta^\frac{\alpha}{K_{jt}} L_{jt}^* = \theta^\frac{\alpha}{K_{jt}} L_{jt}^*.$$  

(36)

The claim follows from (35) and (31):

$$\frac{Y_{jt}}{L_{jt}^*} = \frac{Y_{jt}}{L_{jt}^*} \Leftrightarrow \theta^\frac{\alpha}{K_{jt}} L_{jt}^* = \theta^\frac{\alpha}{K_{jt}} L_{jt}^*.$$  

The last step of the proof is to solve for the relative vintage prices that satisfy the no-arbitrage
condition equating income per dollar of capital across all vintages,
\[
\frac{MPK_{jt}}{P_{jt}} = M_t. \tag{37}
\]
The capital of the latest vintage has price \( P_{jt} = 1 \) with \( J = J_t \), equal to the marginal cost of transforming output into capital. Prices \( P_{jt} \) are determined from evaluating the ratio of relative marginal products of one unit between vintages \( j \) and \( J \) and using (35) and (36):
\[
P_{jt} = \frac{P_{jt}}{K_{jt}} = \frac{MPK_{jt}}{MPK_{jt}} = \frac{Y_{jt} K_{jt}^*}{K_{jt} Y_{jt}} = \frac{Y_{jt}/L_{jt}^*}{K_{jt}/L_{jt}^*} = \frac{K_{jt}^*/L_{jt}^*}{Y_{jt}/L_{jt}^*} = \theta_{j-J} = b^{j-J}.
\]

**Proof of Proposition 2:** Evaluate aggregate output in (1) and the aggregate value of capital in (2) using (36) and \( \sum_{j \leq J} L_{jt}^* = 1 \):
\[
Y_t = \sum_{j \leq J} \theta_0 \cdot [\theta_j^j] [K_{jt}]^\alpha [L_{jt}^*]^{1-\alpha} = \sum_{j \leq J} \theta_0 \cdot [\theta_j^j K_{jt}]^\alpha [L_{jt}^*]^{1-\alpha}
\]
\[
= \sum_{j \leq J} \theta_0 \cdot \left[ \frac{\theta_j^j K_{jt}}{L_{jt}^*} \right]^\alpha L_{jt}^* = \theta_0 \cdot \left[ \frac{\theta_j^j K_{jt}}{L_{jt}^*} \right]^\alpha \sum_{j \leq J} L_{jt}^* = \theta_0 \cdot \theta_J^J \cdot \left[ \frac{K_{jt}}{L_{jt}^*} \right]^\alpha.
\]

\[
K_t = \sum_{j \leq J} P_{jt} \cdot K_{jt} = \theta^\frac{J}{\alpha} \sum_{j \leq J} \theta_j^j K_{jt} = \theta^\frac{J}{\alpha} \sum_{j \leq J} \frac{\theta_j^j K_{jt} L_{jt}^*}{L_{jt}^*}
\]
\[
= \frac{K_{jt}}{L_{jt}^*} \sum_{j \leq J} L_{jt}^* = \frac{K_{jt}}{L_{jt}^*}.
\]

Combining the two expressions above establishes (3).

To establish (4), use no-arbitrage condition (37) setting \( j = J \), and (3):
\[
M_t - \delta = MPK_{jt} - \delta = \alpha \theta_0 \cdot \theta_J^J \left[ \frac{K_{jt}}{L_{jt}^*} \right]^{\alpha-1} - \delta = \alpha \frac{Y_t}{K_t} - \delta.
\]

Lastly, (5) follows directly from (34) in Step 1 of Proposition 1.

**Proof of Proposition 3:** First, show that the limit function \( \Phi^* (K) = \lim_{n \to \infty} \Phi^n (K) \) exists, and it is continuous and non-decreasing. The functions \( \Phi^n \) are continuous and bounded on \((0, K]\). Thus, the Cauchy sequence has a limit \( \Phi^* (K) \), and it is continuous, bounded and non-decreasing.
Next, we prove

**Claim:** There is an interval $\mathbb{K} = [K, \bar{K}]$, with $0 < K < \bar{K} < \infty$ such that all of the rational expectations equilibrium aggregate state realizations are contained in $\mathbb{K}$. For any trajectory originating in $\mathbb{K}$ and obeying the aggregate law of motion (10) — with $\Phi(K) = \Phi^*(K)$ — gross investment is non-negative for all $t \geq 0$.

**Proof** Let the upper bound of $\mathbb{K}$ be the steady state $\bar{K} = \lim_{n \to \infty} K^n$. Checking the upper bound for inclusion of $A_{t+1}$ and $bA_{t+1}/Z$ in $\mathbb{K}$ is straightforward: $K_t \in \mathbb{K} \Rightarrow K_t \leq \bar{K}$. Movement along the saddle path implies that $K_t \leq A_{t+1} \leq \bar{K}$. It follows that $bA_{t+1}/Z \leq \bar{K}$.

Now construct the lower bound, $\underline{K}$. Define

$$
\psi_0(A) = \rho (1 - \lambda) (1 + m(A))
$$

and

$$
\psi_1(A) = \rho \lambda bZ^\gamma - 1 \left(1 + m\left(\frac{bA}{Z}\right)\right).
$$

**Step 1** Show that equations $\psi_0(A) = 1$ and $\psi_1(A) = 1$ each have a unique, strictly positive root, $\underline{A}_0$ and $\underline{A}_1$, respectively. Function $m(A)$ defined in (7) is continuous, strictly decreasing, with $\lim_{A \to 0} m(A) \to \infty$ and $\lim_{A \to \infty} m(A) = -\delta \in (-1, 0)$. Therefore, $\psi_0(A)$ and $\psi_1(A)$ are both continuous, strictly decreasing, with $\lim_{A \to 0} \psi_0(A) \to \infty$, $\lim_{A \to 0} \psi_1(A) \to \infty$ and

$$
\lim_{A \to \infty} \psi_0(A) = \rho (1 - \lambda) (1 - \delta) < 1,
$$

$$
\lim_{A \to \infty} \psi_1(A) = \rho \lambda bZ^\gamma - 1 (1 - \delta) < 1.
$$

We thus established that $\psi_0(A) = 1$ and $\psi_1(A) = 1$ each have a unique, strictly positive root, $\underline{A}_0$ and $\underline{A}_1$, respectively. By construction, $\underline{A}_0$ is the discrete-time analog of, $\bar{K}$, the horizontal intercept of the $\hat{X} = 0$ isocline on Figure 2.

Using the notation $\psi_0(\cdot)$ and $\psi_1(\cdot)$ in Euler equation (18) evaluated on the equilibrium trajectory, we have

$$
u'(\Phi^*(K_t)) = \psi_0(A_{t+1}) \cdot u'(\Phi^*(A_{t+1})) + \psi_1(A_{t+1}) \cdot u'(\Phi^*(\frac{bA_{t+1}}{Z})).$$

(38)

**Step 2** Set $\underline{K} = b/Z \cdot \min\{ \underline{A}_0, \underline{A}_1 \}$. Since $\underline{K} \leq b/Z \underline{A}_0$ and $\underline{A}_0 < \bar{K}$, the interval $\mathbb{K}$ is non-empty. Now show that for any $K_t \geq \underline{K}$, $bA_{t+1}/Z \geq \underline{K}$.

**Case 1:** $K_t \geq \underline{A}_1$. $K_t \in \mathbb{K} \Rightarrow K_t \leq \bar{K}$. Motion on a saddle path requires $A_{t+1} \geq K_t$. Then $bA_{t+1}/Z \geq bK_t/Z \geq b\underline{A}_1/Z \geq \underline{K}$.

**Case 2:** $\underline{A}_1 > K_t \geq \underline{K}$. The proof is by contradiction. Assume that $K_t \geq \underline{K}$ and $bA_{t+1}/Z < K_t < \underline{A}_1$. Since $bA_{t+1}/Z < K_t < \underline{A}_1$, we must have $\psi_1(A_{t+1}) > 1$. From the Euler equation and
monotonicity of $\Phi^*$, we have
\[
u' (\Phi^* (K_t)) > \psi_1 (A_{t+1}) u' \left( \Phi^* \left( \frac{bA_{t+1}}{Z} \right) \right) = \nu' (\Phi^* (K_t)) > u' \left( \Phi^* \left( \frac{bA_{t+1}}{Z} \right) \right)
\]
\[
\Rightarrow \Phi^* (K_t) < \Phi^* \left( \frac{bA_{t+1}}{Z} \right) \Rightarrow K_t \leq \frac{bA_{t+1}}{Z},
\]
a contradiction. Hence, the supposition cannot be correct, and we have established that
\[
A_1 > K_t \geq \frac{bA_{t+1}}{Z} \Rightarrow K_t \geq K.
\]
It follows that $A_{t+1} \geq K$.

Finally, motion on a saddle path establishes that investment is non-negative. ■

**Proof of Proposition 3 continued** We have shown that all of the rational expectations equilibrium aggregate state realizations are contained in a compact interval $K$ that is bounded away from 0. Since K is the feasible set for the household problem (13) and the payoff function $U$ in (12) is strictly increasing in $a_t$, strictly concave in $a_t$ and $a_{t+1}$ and continuously differentiable, the household problem satisfies the assumptions of Theorem 4.15 in Stokey and Lucas (1989). Applying the theorem establishes that the first-order condition and the transversality condition
\[
\lim_{t \to \infty} \sum_{z^t \in Z^t} p(z^t) \rho^t \partial U (K[z^{t-1}], K[z^t], z^t) K[z^t] = 0, \text{ all } z^t \in Z^t. \tag{39}
\]
are sufficient for optimality in (13).

To see that the first-order condition for (13) coincides with (38), collect the terms in (13) that depend on $a_{t+1} = a[z^t]$, differentiate with respect to $a_{t+1}$ and evaluate the resulting expression at $x_t = X_t = \Phi^* (K_t)$, $k_t = K_t$ and $a_t = A_t$.

It is left to show that the rational expectations equilibrium trajectory satisfies the transversality condition (39). This involves two preliminary steps.

**Step 1** Show that
\[
\sigma_t = \sum_{z^t \in Z^t} p(z^t) \rho^t Z^t < 1 \text{ all } t \geq 1.
\]
Since $z^t$ is a sequence of realizations of $t + 1$ i.i.d. Bernoulli random variables, we have
\[
\sigma_t = \sum_{z^t \in Z^t} \lambda^t (1 - \lambda)^{t-J} \rho^t Z^t = \sum_{z^t \in Z^t} (\rho \lambda Z^t)^J \cdot (\rho (1 - \lambda))^{t-J} = \sum_{J=0}^{t} \binom{t}{J} \cdot (\rho \lambda Z^t)^J \cdot (\rho (1 - \lambda))^{t-J} \text{ all } t \geq 1.
\]
Thus, (11) implies that
\[
\lim_{t \to \infty} \sigma_t = 0. \tag{40}
\]

**Step 2** Construct an upper bound on
\[
T_t = \sum_{z^t \in Z^t} p(z^t) \rho^t \frac{\partial U(K[z^{t-1}], K[z^t], z^t)}{\partial a_t} K[z^t]
\]
in (39).

Lemma 1 establishes that \( a[z^{t-1}] \omega(z_t) = K[z^t] \in [K, \bar{K}] \). Use this result to construct an upper bound on the household’s current value of the state:
\[
\frac{\partial U}{\partial a_t} a_t = u'(c_t) \cdot Z^t_k \cdot a_t \cdot (1 + m(K[z^t]))
\]
\[
\leq u'(x_t) \cdot Z^t_k \cdot \bar{K} \cdot (1 + m(K[z^t]))
\]
\[
\leq u'(\Phi^*(\bar{K})) \cdot Z^t_k \cdot \bar{K} \cdot (1 + m(K[z^t])).
\]

Now construct an upper bound on \( T_t \) using the above expression:
\[
T_t \leq u'(\Phi^*(\bar{K})) \cdot \bar{K} \cdot (1 + m(K[z^t])) \cdot \sigma_t.
\]

Then (40) establishes (39).