Conditional Stochastic Disasters and the Equity Premium Puzzle

Alex Chen
Department of Economics
Stanford University
alechen@stanford.edu

May 2012

Under the direction of
Professor John Shoven & Professor Kenneth Judd

Abstract
Motivated by Barro’s (2006) approach towards modeling the equity premium, we study the original formulation of the equity premium puzzle as posed by Mehra and Prescott (1985). Firstly, we build a single-regime stochastic model capable of simulating the effects of rare economic disasters. We then extend that into a multi-regime model which allows study of conditional disaster regimes and shocks that are not i.i.d., allowing for the simulation of a robust variety of economic phenomenon. Finally, we examine the results of modeling depression-recession economic cycles with the possibility of recessionary aftershocks following severe depressions. Compared to historical economic data, we conclude that a multi-regime model can account for the observed equity premium while simultaneously producing a plausibly low real rate.

Acknowledgments: I am immensely grateful to my advisors, Professor John Shoven and Professor Kenneth Judd. Their knowledge, expertise, and advice all proved invaluable in the creation of this thesis. I could not have completed this work without their economic insights and mathematical guidance. I am appreciative of comments from Professor Geoffrey Rothwell, Nipun Kant, Sijia Wang, and my father, Professor Hongwei Chen. Thank you to my family for their support and love.
Contents

1 Introduction 4

2 Motivation 5

3 Literature Review 8
   3.1 Mehra and Prescott’s Original Design 9
   3.2 Further Studies into the Existence of the Equity Premium Puzzle 12
   3.3 Rietz’s Disaster Theory Extension of Mehra and Prescott’s Model 13
   3.4 Barro’s Revival of the Disaster Theory 15

4 Building a Computational Model 17
   4.1 Equity Returns 19
   4.2 Riskless Asset Returns 21
   4.3 The Single-State Case 21
   4.4 A Two-State Model 23
   4.5 Generalizing to n-states 27
   4.6 Determining Average Equity Premia 28
   4.7 Reproducing Mehra and Prescott’s Original Results 30

5 Implementation of Stochastic Matrices 30
   5.1 Modeling Standard Variability of Consumption Growth 31
      5.1.1 The Simple Symmetric Random Walk 31
      5.1.2 The Gaussian Random Walk 32
      5.1.3 Building Mean Reversion 35
   5.2 Modeling Rare Economic Disasters 37

6 Incorporating Stochastic Regimes 39
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Equity Returns Under Multiple Regimes</td>
<td>40</td>
</tr>
<tr>
<td>6.2</td>
<td>Riskless Asset Returns Under Multiple Regimes</td>
<td>43</td>
</tr>
<tr>
<td>6.3</td>
<td>Average Equity Premia Under Multiple Regimes</td>
<td>43</td>
</tr>
<tr>
<td>7</td>
<td>Conditional Stochastic Disasters</td>
<td>44</td>
</tr>
<tr>
<td>7.1</td>
<td>A Conditional Multi-Regime Model</td>
<td>45</td>
</tr>
<tr>
<td>8</td>
<td>Concluding Remarks</td>
<td>46</td>
</tr>
<tr>
<td>9</td>
<td>References</td>
<td>48</td>
</tr>
<tr>
<td>A</td>
<td>Tables of Results</td>
<td>50</td>
</tr>
<tr>
<td>B</td>
<td>Stochastic Matrices</td>
<td>52</td>
</tr>
<tr>
<td>B.1</td>
<td>( \Pi_{SS} )</td>
<td>52</td>
</tr>
<tr>
<td>B.2</td>
<td>( \Pi_{G} )</td>
<td>52</td>
</tr>
<tr>
<td>B.3</td>
<td>( \Pi_{MR} )</td>
<td>53</td>
</tr>
<tr>
<td>C</td>
<td>Determination of Model Variables</td>
<td>54</td>
</tr>
</tbody>
</table>
1 Introduction

Conventional economic and financial wisdom tells us that equity is a riskier investment than government debt. Consequently, equity necessarily produces much higher returns in order to compensate for the greater risk of holding it. Indeed, a 1926 investment of $1.00 in U.S. Treasury bills (short-term government debt, generally considered to be one of the safest securities in the world) would have yielded a real value of $1.70 at the end of 2009, for an average annualized inflation-adjusted return of 0.63%. The same investment in the S&P 500, a diversified index of large U.S. company stocks, would have returned a real value of $215.15, an average annualized inflation-adjusted return of 6.6%.

Such a spectacular difference in returns over this 84-year period stems from an average annual premium of 6% on equity returns over those of the safer Treasury bills, and underscores the massive importance of this premium to portfolio allocation and investment decisions.

Historically observed results concur with established economic theory: stocks are indeed riskier investments than Treasury bills due to larger uncertainty in stock returns. As rational investors look to maximize returns while minimizing risk, stocks must command a higher expected average return in order to convince investors to hold them. In the absence of an equity premium (in other words, if stocks and government bonds have the same returns) but the same variances in returns as historically observed, there would be no demand from investors for equity as they could generate the same returns with less risk by investing solely in the safe asset. Thus, stocks must, on average, have higher returns than bonds in order to compensate for the greater variance in their returns.

This leads to an analogous question: when are bond returns low enough that investors no longer have the motivation to hold them over stocks? At some point, returns on equity

\(^1\)Data from Ibbotson SBBI 2010, pg. 25-27.
become high enough that they ought to more than compensate for the increased variance of returns. Just as stocks generating about the same returns in the long-run as bonds results in a lack of equilibrium in the market, there exists a point where stock returns are high enough that investing in bonds no longer makes sense.

Mehra and Prescott (1985) set out to determine what the proper equity premium should be, given historical observations in the U.S. stock market from 1889 to 1978. They concluded that the observed average annual U.S. equity premium of 6% completely contradicted economic models’ predictions: only a 0.35% equity premium could be properly justified in the scope of their models. Equity simply does not possess sufficiently more risk than Treasury bills to produce the high observed equity premium; alternatively individuals are simply not risk-averse enough to be indifferent between the average rates observed on safe assets and the average observed returns and variance on equity. Mehra and Prescott dubbed this enigma the ‘equity premium puzzle.’

The equity premium puzzle has proven robustly resistant to economic academia’s attempts to resolve it. Twenty-seven years after the initial proposal of the puzzle, no widely accepted answer yet exists.

2 Motivation

Fundamentally, stocks ought to return a premium over Treasury bills based on the higher relative variance in stocks’ returns. However, merely considering the relative variances of returns for stocks and bonds does not properly encapsulate their differences in risk. When investing in stocks, investors are wary of the potential for sudden, severe decreases in the value of equity as a possible risk associated with such investments. Similarly, bonds (even the near-riskless T-bills) have some degree of counterparty default risk associated with them. Making considerations for the structure of these financial instruments in our model (as
investors certainly do) allows us to better model the equity premium.

As a second matter of consideration, events such as economic depressions and recessions severely contract the value of equity, and even though the impact of such events is dampened when considering average returns over long-term historical periods or over periods in which a depression happened to not have occurred, accounting for them potentially produces a much more reasonable theoretical equity premium. Empirically, investor consideration for stock market crashes can be seen in the prices of equity options: deep out-of-the-money downside strikes tend to have higher implied volatilities (thus commanding a premium due to the higher volatility) than the at-the-money and in-the-money strikes. Out-of-the-money put options, which severely appreciate in value with steep drops in the prices of their underlying security and hence can be considered insurance against stock market crashes, trade at a premium for the level of risk associated with them compared to other options on the same security. This phenomenon, known as the volatility skew, demonstrates that the market demands a premium to hedging against the potential for sudden, severe drops in equity price, regardless of whether such a risk is actually likely. Studies have found that information on market crashes can be derived from the prices in the options market (Dorran, Tarrant, and Peterson 2007). Such evidence in observed financial security prices motivates the consideration of low-probability economic depressions. The size and direction of stock price fluctuations do not exactly follow a log-normal distribution, thus ignoring the long tail of low-probability events has a non-negligible effect on our results.

In their original model, Mehra and Prescott (1985) only considered the differences in variance of returns when weighing the relative riskiness of stocks and of bonds. Rietz (1988) first introduced the possibility of low-probability economic disasters as a potential solution, modifying Mehra and Prescott’s original model to incorporate the possibility of severe contractions in consumption (akin to a stock market crash). Barro (2006) further refined Rietz’s idea of incorporating low-probability depressions and built a more robust model which also
accounted for the potential for partial default of government debt. Importantly, Barro undertook a detailed analysis of economic depressions across modern global history, finding that the disaster scenarios originally envisioned by Rietz are actually in the realm of reality, in contrast to the dismissal of them as improbably severe by Mehra and Prescott (1988). Further discussion of Rietz and Barro’s ideas and work can be found in the literature review.

Rietz and Barro’s results concluded that consideration of uncommon depressions which severely contract the value of equity justify a considerably larger equity premium compared to the results of Mehra and Prescott’s original study. An implausibly high level of risk aversion is no longer necessary to produce a 6% equity premium in the model once we consider the potential for sudden and severe tumbles in stock prices. We look to further improve the simulation of the underlying market structure by changing the manner in which disasters occur in our model. Rather than implementing an independent stochastic term which contracts consumption by a fixed percentage with a predetermined probability of occurring in each period, we model economic depressions as systemic and dependent on past history.

Consider that there exists some low probability that an economic depression will occur during a given year. The value of this probability depends on the state and health of the economy, hence it would be higher immediately following an economic depression compared to in times of prosperity. Barro’s model treats the disaster probability as a constant (determined by looking at the averages in 20th century historical data from the 20 founding OECD countries) when in fact a more realistic treatment should use a variable disaster probability conditional on the state of the economy and/or the length of time since the previous depression.

What happens following an economic depression? Output shrinks, unemployment spikes upwards, uncertainty in the economy increases, and equity prices plummet. As a consequence, real GDP per capita sharply falls, and we see a sustained period of actual output
severely lagging potential output. During and soon after the occurrence of the depression, the recovery period begins and growth increases to higher levels than normal. However, when the economy has not returned to full strength, it is plausible to think that increased risk of economic aftershocks exists. Consider the Great Depression, which most economists refer to the period of 1929-1941. The National Bureau of Economic Research, which publishes business cycle dates for the U.S. economy, notes that a second recessionary period occurred from 1937 to 1938. Rather than one long depression, the 1929-1941 period should actually be thought of as a serious economic disaster (starting with the original 1929 stock market crash) and a period of recovery, interrupted by a smaller recessionary aftershock in 1937.

With such considerations in mind, we refine the insights of Rietz and Barro and establish a model which has multiple stochastic regimes, representing different degrees of economic health. In essence, this is similar to Rietz’s three-state Markov chain specification, but the purpose of the multiple regimes in our model is not to capture normal noise in economic growth, as it is in Rietz’s model. Rather, each regime has its own different state transition probabilities, and our model economy moves between these regimes according to a stochastic process. In each period consumption changes according to a regime-specific stochastic model, which accounts for consumption growth, variance in consumption growth, and the potential for rare, severe depressions. With multiple regimes and states, we simulate sophisticated economic cycles and study their effects on the equity premium.

3 Literature Review

Here, we summarize the literature relevant to the ideas pursued in this paper. Mehra and Prescott’s (1985) model design and assumptions are central. We also describe Rietz (1988) and Barro’s (2006) respective implementations and studies of low-probability eco-

\[2\text{http://www.nber.org/cycles/cyclesmain.html} \]
nomic crashes in great depth.

Prior to 1978, questions on asset pricing and returns were primarily studied within the framework of the capital asset pricing model (CAPM), built on work by Markowitz (1952), and Sharpe (1964). Because CAPM treats the equity risk premium as an input rather than a product of the model’s results, it was unsuitable for direct study on the nature of the equity premium itself. Lucas (1978) provided a watershed theoretical pure exchange model that encouraged a much deeper examination of equilibrium asset prices. His model, often-referred to as the Lucas tree model, becomes the dominant framework for economic literature on asset pricing and the equity premium.

3.1 Mehra and Prescott’s Original Design

Mehra and Prescott (1985) set the important precedent of pioneering an asset pricing model capable of simulating the equity premium, and then calibrating it to match the observed characteristics of the U.S. economy. They employ an adaptation of the pure exchange Lucas tree model (1978) with a single representative consumption good and stationary equilibrium growth of consumption and asset returns.

Mehra and Prescott’s model assumes a Markov process for the growth rate of the economy’s endowment.\(^3\) Such a design choice captures movement in consumption caused by large changes in GDP per capita, such as the many-fold increase in GDP per capita seen in the U.S. over the past 100 years. The economy in Mehra and Prescott’s model consists of a single representative household with preferences determined by \(E_0 \{ \sum_{t=0}^{\infty} \beta^t u(c_t) \} \), where \(c_t\) is per capita consumption and \(\beta\) is the discount factor over each successive time period. The functional form denotes the expectation operator conditional upon information available at the current period \((t = 0)\). Mehra and Prescott choose the constant relative risk aversion

\(^3\)This is in contrast to Lucas’s version of the model which employs a Markov process for the level of endowment, rather than its growth rate.
(CRRA) utility function \( u(c, \theta) = \frac{1}{1-\theta} \), with \( 0 < \theta < \infty \) representing the degree of curvature of the utility function. When \( \theta = 1 \), the utility function is logarithmic.

Mehra and Prescott make the additional assumption (maintained from Lucas’s model) that there is only one productive unit and one competitively traded equity share entitling the owner to the production yield of the tree. Since there is only one productive unit, the return \( y_t \) on the equity share is equivalent to the return on the market for that period. Their Markov process setup designates an endowment growth rate of \( y_{t+1} = x_{t+1}y_t \), where \( x_{t+1} \in \{\lambda_1, \lambda_2\} \) is the growth rate between the current period and the next and is determined by an ergodic Markov chain with two assumed states. One state represents high growth with growth rate \( \lambda_1 = 1 + \mu + \delta \) while the other represents low growth with growth rate \( \lambda_2 = 1 + \mu - \delta \), which models variable growth either above or below the mean trend. Furthermore, they define a probability transition matrix of the form

\[
\begin{pmatrix}
\Pi_{11} & \Pi_{12} \\
\Pi_{21} & \Pi_{22}
\end{pmatrix}
= \begin{pmatrix}
\Pi & 1 - \Pi \\
1 - \Pi & \Pi
\end{pmatrix},
\]

where \( 0 < \Pi_{ij} < 1 \) is the probability of being in the state \( j \) of the Markov chain in the next period, given being in state \( i \) in the current period.

Such a specification of the Markov chain allows independent variation of the model’s average growth rate through control of \( \mu \), variability of growth through control of \( \delta \), and the serial correlation of growth rates through adjustment of \( \phi \). Studying sample values for the U.S. economy for the period 1889-1978, Mehra and Prescott used parameters \( \mu = 0.018 \), \( \delta = 0.036 \), and \( \Pi = 0.43 \).

Using their model, Mehra and Prescott derive expressions for the equilibrium price of the equity share and risk-free bill. They then measure the predicted equity premium within their model with estimates for the parameters derived from U.S. historical data. However, the largest equity premium that the authors manage to obtain with their model is 0.35%, which
strongly conflicts with the observed equity premium of over 6%. Hence, Mehra and Prescott (1985) describe the equity premium puzzle as the unexplainably higher equity premium observed in historical data compared to what economic models are capable of justifying.

As we see with later work in the literature, the design choices originally made by Mehra and Prescott prove very influential. Their model ensures a stationary equilibrium return process as well as uses a representative agent whose preferences exist independent of initial endowment distribution.

The paradox that Mehra and Prescott find in their study is that with real consumption per capita rising at a rate of approximately 2% per year (the approximate observed economic growth rate of the U.S. and other developed nations), the elasticity of substitution between consumption in the present year $t$ and the next year $t + 1$ must be relatively low. Such a value for the elasticity of intertemporal substitution cannot justify an equity premium anywhere near as high as 6%. More precisely, for the equity returns to approximately average the historically observed 7%, the real rate of return (equivalent to the return on U.S. government bonds) should be much higher than the observed 1%.\(^4\) Because the economy is growing on average and thus individuals expect future consumption to exceed consumption in the present, the marginal utility of consumption in the present surpasses marginal utility in the future. Thus, interest rates on average must be higher and much closer to the return on equity.

Mehra and Prescott note that the neoclassical economic growth model cannot possibly explain the historical equity premium of 6%. They show that such a large equity risk premium, given observed consumption growth in the United States over the same time period, can only be reconciled with an implausible level of risk aversion in order to accord with the standard model. They also state that their conclusions are robust, with changes in the average growth rate of consumption between 1.4% and 2.2% having little effect on

\(^4\)This describes the low risk-free rate problem, an analogous asset pricing puzzle.
the model’s maximum predicted equity premium of 0.35%. Thus, they suggest a rejection of Arrow-Debreu competitive equilibrium models in order to rationalize the excessively high equity risk premium observed.

### 3.2 Further Studies into the Existence of the Equity Premium Puzzle

As an alternative to rejecting Arrow-Debreu competitive equilibrium models, economists have comprehensively examined economic data from other capital markets and historical periods. Additional studies on the equity premium have supported the existence of the puzzle. Mehra (2003) points out that U.S. data on stock and bond returns from the past 150 years and equity returns in the capital markets of other developed countries (such as France, Germany, and Japan) also exhibit a similar equity premium.

Cecchetti, Lam, and Mark (1993) take the view that the equity premium puzzle is actually a natural statement of preferences and that the standard literature underestimates investors’ degree of risk aversion. Their argument asserts that increasing estimates of the degree of risk aversion better justifies the 6% premium seen in the spread between stock returns and T-bill returns. However, even extreme values of the relative risk aversion coefficient $\theta$ (for instance $\theta = 10$) fail to resolve Mehra and Prescott’s original formulation of the puzzle, as raising $\theta$ to a level capable of producing the observed historical equity premium also raise the model’s real rate to over 6%, which also clashes with historical data. Furthermore, most economists believe that such high $\theta$ implies extremely implausible behavior on the part of individuals (Kocherlakota 1996).

Kurz (2001) suggests that asset allocation puzzles such as the equity premium puzzle are

---

5For example, an individual with $\theta = 10$ would be indifferent between a guaranteed payout of $53,991 and a bet equally likely to pay $50,000 or $100,000. If $\theta = 20$, the necessary value of the guaranteed payoff for indifference falls further to $51,858 (Mankiw and Zeldes 1991). Hence, most economists deem $\theta > 5$ to be implausibly high.
simply indicators of the untenability of the rational expectations hypothesis upon which the neoclassical economic growth model is grounded. He expresses a measure of surprise that the robust and extensive literature attempting (and thus far failing) to satisfactorily solve the equity premium puzzle has not yet led to a rejection of rational expectations.

As an alternative, Kurz proposes the theory of rational beliefs which provides a resolution to the asset allocation puzzles that appear under rational expectations. Under rational beliefs, economic agents do not possess complete structural information about the market, and thus cannot make an optimal guess of the true valuation of an asset. Kurz takes the view that price variability attributed to the beliefs of economics agents, which he refers to as endogenous uncertainty, “is the dominant form of uncertainty in the equities markets.” Since Mehra and Prescott (1985) calculate an expected equity premium under rational expectations, ignoring endogenous uncertainty, Kurz argues that their model does not account for a significant amount of uncertainty present in equity. Equity holders require additional compensation for taking on this additional risk, which justifies a much greater equity premium than can be predicted under rational expectations.

Most economists are reluctant to accept Kurz’s argument, as it requires abandoning standard preferences and the rational expectations hypothesis. We adhere to rational expectations and the neoclassical growth model, as we aim to reach a rational and credible solution to the equity premium puzzle within the scope of the original framework of the equity premium puzzle as stated by Mehra and Prescott.

3.3 Rietz’s Disaster Theory Extension of Mehra and Prescott’s Model

Rietz (1988) proposes one of the earliest solutions to Mehra and Prescott’s challenge by suggesting the existence of low-probability market crashes such as the Great Depression as a
method of resolving the disparity between the predicted and observed equity premium. Rietz modifies Mehra and Prescott’s (1985) original specification of a two-state Markov process to model more extreme consumption variation across periods. Rietz suggests a three-state specification, adding a rare depression-like state to the original two possible states of an above-average and below-average economy. An economic crash resulting in the third state (which occurs following a good or bad period with some constant low probability $p$) causes consumption to fall some to fraction $(1-b)y_t$ for that period (with $0 < b < 1$ representing the percentage that consumption decreases by during a crash). The consumption growth rate returns to normal in the next period, but the consumption level is permanently affected, since the Markov process dictates growth rates rather than endowment levels.

Using his model, Rietz calibrates parameter values for crash probability, the relative risk aversion parameter $\theta$, and decrease in output in the case of a market crash. Using values of contraction fraction $b = 0.25$ or $0.5$ (output falling by one-quarter or one-half of its normal expected value in a crash), Rietz produces reasonable parameters estimates for which the model predicts similar risk-free rates and risk premium to those observed in the empirical data. However, his necessary values for the risk aversion parameter $\theta$ remain much higher than are generally accepted. He goes on to model the hypothetical possibility of an extreme economic crash which destroys a period’s entire endowment other than the approximately 2% from economic growth, equivalent to $b = 0.98$. Incorporating the possibility of such a crash (with a crash probability $p < 0.1\%$) allows the risk aversion parameter value to be less than two, and thus in the plausible range predicted by empirical studies (Mehra and Prescott 1985). By altering the original model in such a fashion, Rietz manages to produce an equity premium similar to that observed in historical data while adhering to an Arrow-Debreu type asset pricing model, thus allowing for a Debreu competitive equilibrium with non-stationary

\footnote{Mehra and Prescott also experiment with a four-state Markov model (good, bad, average good, and average bad) but find that it yields a risk premium of only 0.39%; virtually identical to the low 0.35% predicted by their original specification.}
consumption levels.

Mehra and Prescott (1988) were quick to respond to and reject Rietz’s proposed theory. They argue that the disaster scenarios proposed by Rietz (with an one-year decline in consumption between 25% and 98%) are far too severe and historically unprecedented. They also note that equating the return on Treasury bills with the real rate is not feasible during times of economic disaster due to high and sudden unanticipated inflation (which can decrease the return on Treasury bills and even possibly renders them completely worthless in the most extreme scenarios). They also argue that the value of the risk aversion parameter has been empirically estimated to be approximately $\theta = 1$, but Rietz’s model requires a value of at least 5 and often near 10 to produce an equity premium close to the observed 6%. As a final counterargument, Mehra and Prescott point to interest rate movements in times of crisis or war as counter to what Rietz’s model predicts.

3.4 Barro’s Revival of the Disaster Theory

Yet recent research has revived Rietz’s disaster theory as a plausible way to resolve the equity premium puzzle with standard economic preferences. Weitzmann (2007), using Bayesian estimation of uncertain structural growth, suggests that “parameter uncertainty...adds a permanent tail-thickening effect to posterior expectations”, akin to the low-probability disaster scenario. Thus, the fear of a market crash (whether justified or not) is enough to drive up the equity premium by producing uncertainty among investors. Weitzmann’s Bayesian argument parallels Kurz’s idea (2001) of building investor uncertainty in the market into an economic hypothesis. Motivated by such positive results, Barro (2006) reexamines building the rare but devastating economic disaster into an economic model studying the size of the equity premium. His model integrates Rietz’s disaster scenarios by employing a design with stochastic consumption growth rate modeled as a random walk with drift. Barro’s model assumes a known constant $p \geq 0$ to be the probability of a disaster occurring in a period,
with output contracting by a random variable (rather than taking Rietz’s approach of using a predetermined constant factor, such as 0.25 or 0.5) in the case of such an economic disaster.

Barro also uses the distribution of 20th century economic data to calibrate his model’s constants, notably the value of the disaster probability $p$. Furthermore, in his study of international economic data of the past century, Barro points out that disaster scenarios of the magnitude that Rietz suggested (where consumption falls by 50% or more in a period) actually are not historically unprecedented.\(^7\) Mehra, Prescott, and other economists previously viewed such severe crashes as unrealistic and purely theoretical. Barro argues that $p = 1.5\% - 2\%$ and disaster sizes $0.15 < b < 0.64$ are reasonable model values given the distribution of 20th century historical data.

Importantly, Barro also adds firm leverage to the model. By considering three types of assets - equity, government bonds, and private debt - he shows that a pure exchange model with just equity and government bonds needs to produce an equity premium that is only $\frac{2}{3}$ the size of that historically observed, as adding leverage typical of the American corporate sector increases the model equity premium by a factor of 1.5. Such a multiplier is insignificant with an equity premium in the neighborhood of 0.3%, as Mehra and Prescott calculated it to be, but with the incorporation of rare disasters or other factors attempting to resolve the equity premium puzzle, the effect of firm leverage becomes substantial.

In a review of the literature regarding the equity premium puzzle over the two decades since he and Prescott first proposed it, Mehra (2006) states that Barro’s results “are intriguing and lend support to a resurrection of the disaster scenario as a viable justification for the premium.”

\(^7\)Barro explains that every OECD country except Switzerland experienced a 15% or greater decline in real GDP per capita in the twentieth century. Germany experienced a 64% fall in real GDP per capita during the 1944-46 period.
4 Building a Computational Model

In order to study the consequences of the conditional economic depressions and different market states that we propose, we build a theoretical construction which simulates the evolution of returns on assets as our model economy grows through time.

Similarly to Mehra and Prescott (1985), we employ a variant of the Lucas tree model (1978) for asset pricing. Our model economy has a single consumption good and a single productive unit (the metaphorical ‘tree’) which produces a dividend of the good in each period. Our consumption good is perishable and no storage mechanisms exist to transfer dividends into future periods, so each period’s consumption in our model equals that period’s dividend. The length of a period can be arbitrary, although in this paper we typically use one year as the period length unless otherwise stated.

There exists a single equity share on the productive unit, entitling the holder of the equity share to the dividend produced in that period. There also exists a risk-free asset which guarantees a riskless return, and can be thought of as the equivalent of a Treasury bill.

Our model economy moves both stochastically through \( n \) different possible states and deterministically forward in time. Each state represents a possible deviation of the economy from its baseline growth trend. The stochastic variation in states models uncertainty in economic growth and returns, as well as the potential for rare, serious depressions. The movement through time allows us to model returns over time while building in baseline economic growth around which the economy varies.

Hence, our model differs from Mehra and Prescott’s 1985 model in that growth levels do not follow a Markov process (and are instead deterministic), but rather that we build stochastic variation into the dividend levels at each period. Hence variance in the growth comes from the state transitions, which are normally distributed. We discuss the stochastic
matrices used in section 5.

Every state $i$ at period $t$ has price $p_{i,t}$ for the equity share, dividend $d_{i,t}$ paid to the owner of that share, and riskless asset with return $r_{i,t}^f$. A single representative consumer exists and seeks to maximize the present value of all current and future consumption over random paths given some utility function $u(c_t)$ and time discount factor $\beta$.

Consequently, the relationship between the price of the equity share and its dividend follows the first-order condition solution of the Lucas model

$$p_{i,t}u'(c_{i,t}) = \beta E_t \left[ u'(c_{t+1})(p_{t+1} + d_{t+1}) \right],$$

while the riskless return in state $i$ at period $t$ obeys the equation

$$u'(c_{i,t}) = \beta E_t \left[ u'(c_{t+1})(1 + r_{i,t}^f) \right] = \beta(1 + r_{i,t}^f) E_t \left[ u'(c_{t+1}) \right].$$

In the above expressions (1) and (2), the expectation operator $E_t$ denotes expectation conditional upon all information known at period $t$.

The $n$ possible states of our model represent a discretization of the possible continuous distribution of dividend yields. For any period, the dividend in any state $j$ is related to the dividend of any other state $i$ by the relation

$$d_{j,t} = \delta^{j-i}d_{i,t},$$

where $\delta$ is a model constant that controls the standard deviation of consumption growth. This ensures that the expected dividend growth process for each state is time-invariant, and the expected dividend for each state for any period is identical to that of any other period with the addition of some exponential multiple of the growth rate.
At the end of each period, the economy moves to a different state via a stochastic process controlled by a stochastic probability transition matrix Π. Furthermore, every state’s dividend increases by growth rate multiple \( G > 1 \) at the end of each period, hence \( d_{i,t+1} = Gd_{i,t}, \forall i, t \). Hence, the economy moves stochastically through these \( n \) possible states of dividend yields which themselves are all rising at a fixed rate. The relation of different states with one another is further detailed in section 5, when we construct our stochastic probability transition matrices to model the types of asset price movements desired.

Our \( n \)-state stochastic specification allows us to rewrite the expectation in (1) as

\[
p_{i,t}u'(c_{i,t}) = \beta \sum_{j=1}^{n} \Pi_{ij} \left[ u'(c_{j,t+1})(p_{j,t+1} + d_{j,t+1}) \right],
\]

(4)

where \( \Pi_{ij} \) is the probability of being in state \( j \) in the next period, conditional on being in state \( i \) in the current period. We control the nature of our model’s stochastic behavior through specification of the probability transition matrix \( \Pi \). Analogously to (4), we can re-express (2) as

\[
u'(c_{i,t}) = \beta (1 + r_{i,t}^{f}) \sum_{j=1}^{n} \Pi_{ij} u'(c_{j,t+1}).
\]

(5)

With this model, we can determine the returns for the equity share and the riskless asset for each period as the economy moves stochastically through the \( n \) states and forward in time, allowing us to calculate each period and potential dividend state’s equity premium.

### 4.1 Equity Returns

Starting with the initial first order condition expressed in (4), we have that \( c_t = d_t \) for all \( t \) in equilibrium (as our model specifies that consumption equals dividends in each period), giving us

\[
p_{i,t}u'(d_{i,t}) = \beta \sum_{j=1}^{n} \Pi_{ij} \left[ u'(d_{j,t+1})(p_{j,t+1} + d_{j,t+1}) \right],
\]

(6)
Following the example of Mehra and Prescott (1985), Barro (2006), and other authors in the literature, we apply the constant relative risk aversion utility function

\[ u(c) = \frac{c^{1-\theta}}{1-\theta}, \quad (7) \]

where \( \theta > 0 \) is the parameter of relative risk aversion and \( u(c) = \log(c) \) when \( \theta = 1 \).

Combining (6) and (7), we have

\[ p_{i,t}d_{i,t}^{-\theta} = \beta \sum_{j=1}^{n} \Pi_{ij} \left[ d_{j,t+1}^{-\theta}(p_{j,t+1} + d_{j,t+1}) \right], \]

First, we divide both sides by \( p_{i,t} \), which gives us

\[ d_{i,t}^{-\theta} = \beta \sum_{j=1}^{n} \Pi_{ij} \left[ \frac{d_{j,t+1}^{-\theta}(p_{j,t+1} + d_{j,t+1})}{p_{i,t}} \right]. \]

Now, we multiply both sides by \( d_{i,t}^{\theta}/\beta \), which gives us

\[ \frac{1}{\beta} = \sum_{j=1}^{n} \Pi_{ij} \left[ \frac{d_{j,t+1}^{-\theta}(p_{j,t+1} + d_{j,t+1})}{p_{i,t}d_{i,t}^{\theta}} \right]. \quad (8) \]

We then use (8) to incorporate growth into our model, giving

\[ \frac{1}{\beta} = \sum_{j=1}^{n} \Pi_{ij} \left[ \left( \frac{d_{j,t+1}^{-\theta}}{d_{i,t}^{-\theta}} \right) \left( \frac{p_{j,t+1} + d_{j,t+1}}{p_{i,t}} \right) \right] = \sum_{j=1}^{n} \Pi_{ij} \left[ G_{ij}^{-\theta} \left( \frac{p_{j,t+1} + d_{j,t+1}}{p_{i,t}} \right) \right], \quad (9) \]

where \( G_{ij} = d_{j,t+1}/d_{i,t} \).
4.2 Riskless Asset Returns

Starting with the equation for the riskless rate (5), we apply \( c_t = d_t \) in equilibrium for all \( t \) and the constant relative risk aversion utility function (7), giving

\[
d_{i,t}^{-\theta} = \beta (1 + r_{i,t}^f) \sum_{j=1}^{n} \Pi_{ij} d_{j,t+1}^{-\theta}.
\]

(10)

Solving for \( R_{i,t}^f = r_{i,t}^f + 1 \), we have

\[
R_{i,t}^f = \frac{d_{i,t}^{-\theta}}{\beta \sum_{j=1}^{n} \Pi_{ij} d_{j,t+1}^{-\theta}}.
\]

(11)

4.3 The Single-State Case

Consider the simplest iteration of our model: an economy with \( n = 1 \) states. Hence, at the end of every period, no stochastic variation occurs in the economy because the only possible state transition is from state 1 to itself. After each period \( t \), the dividend for this state \( d_{1,t} \) increases by growth multiple \( G \).

Without random variation in dividends (and hence in consumption), there is no expectation in consumption because the dividend for the next period (as well as for any future period) is perfectly known from today’s dividend.

Here, we determine the expressions for returns on equity and the riskless asset. Because we have no stochastic variation of states, our probability transition matrix \( \Pi \) is a \((1 \times 1)\) matrix with \( \Pi_{11} = 1 \). Hence, (9) simplifies to

\[
\frac{1}{\beta} = G^{-\theta} \left( \frac{p_{t+1} + d_{t+1}}{p_t} \right),
\]

(12)

where \( G \equiv d_{t+1}/d_t \). Note that we only denote the period subscript here and not the state.
subscripts, as we only have one state in this simplest case.

We are interested in the single-period return on equity $R^e_t$,

$$R^e_t = E_t \left[ \frac{p_{t+1} + d_{t+1}}{p_t} \right], \quad (13)$$

and hence from (12) we have

$$R^e_t = E_t \left[ \frac{p_{t+1} + d_{t+1}}{p_t} \right] = \left( \frac{p_{t+1} + d_{t+1}}{p_t} \right) = \frac{G^\theta}{\beta}. \quad (14)$$

Similarly to equity returns in the single-state case, we see that (11) simplifies to

$$R^f_t = \frac{d_t - \theta}{\beta d_{t+1}} - \frac{\theta}{\beta + 1} = \frac{G^\theta}{\beta}. \quad (15)$$

Combining (14) and (15), we see that the period-specific equity premium is

$$\text{Equity Premium}_t = R^e_t - R^f_t = \frac{G^\theta}{\beta} - \frac{G^\theta}{\beta} = 0. \quad (16)$$

Hence there is no equity premium in our model economy’s single-state case. This result makes intuitive sense, as equity in the single-state case actually functions identically to the riskless asset and does not command a premium. There is no variance or uncertainty in returns, hence $R^e_t$ exactly equals $R^f_t$ for all periods $t$.  

Page 22 of 54
4.4 A Two-State Model

Now, consider a version of the model which has $n = 2$ states: state 1 and state 2. State 1 represents the economy being in a below-trend period while state 2 represents the economy being above-trend. After each period, there is some probability $\Pi_{ii}$ of remaining in the current state $i$ and some probability $\Pi_{ij}$ of moving to the other state $j$.

With our above specification, from (9) we have the following equalities from equity prices

$$\frac{1}{\beta} = \sum_{j=1}^{2} \Pi_{1j} \left[ G_{1j}^{-\theta} \left( \frac{p_{j,t+1} + d_{j,t+1}}{p_{1,t}} \right) \right]$$

$$= \sum_{j=1}^{2} \Pi_{2j} \left[ G_{2j}^{-\theta} \left( \frac{p_{j,t+1} + d_{j,t+1}}{p_{2,t}} \right) \right].$$

From the first equality we have

$$\frac{1}{\beta} = \Pi_{11} \left[ G_{11}^{-\theta} \left( \frac{p_{1,t+1} + d_{1,t+1}}{p_{1,t}} \right) \right] + \Pi_{12} \left[ G_{12}^{-\theta} \left( \frac{p_{2,t+1} + d_{2,t+1}}{p_{1,t}} \right) \right]$$

$$= \Pi_{11} \left[ G_{11}^{-\theta} \left( \frac{p_{1,t+1} + d_{1,t+1}}{p_{1,t}} \right) \right] + \Pi_{12} \left[ G_{12}^{-\theta} \left( \frac{p_{2,t+1} + d_{2,t+1}}{p_{1,t}} \right) \right]$$

$$= \Pi_{11} \left[ G_{11}^{-\theta} \left( \frac{p_{1,t+1} + d_{1,t+1}}{q_{1}d_{1,t}} \right) \right] + \Pi_{12} \left[ G_{12}^{-\theta} \left( \frac{p_{2,t+1} + d_{2,t+1}}{q_{1}d_{1,t}} \right) \right].$$

(17)
where the price-dividend ratio \( q_i = p_{i,t}/d_{i,t}, \forall t \). Similarly, the second equality gives us

\[
\frac{1}{\beta} = \Pi_{21} \left[ G_{21}^{-\vartheta} \left( \frac{q_1 d_{1,t+1} + d_{1,t+1}}{q_2 d_{2,t}} \right) \right] + \\
\Pi_{22} \left[ G_{22}^{-\vartheta} \left( \frac{q_2 d_{2,t+1} + d_{2,t+1}}{q_2 d_{2,t}} \right) \right].
\]

(18)

The equality between (17) and (18) gives us two equations with two unknowns, \( q_1 \) and \( q_2 \). Due to the structure of our model’s dividend payouts, our price-dividend ratios \( q_i \) are time-invariant and depend only on the state. Because ownership of an equity share entitles its owner to all future dividend payments, the price of an equity share equals the value of the infinite discounted sum of all future dividends. At each period, the potential future dividends process for any state is identical to that of any past or future period’s for that same state.

Hence, the price of any equity share and its dividend payments rise at exactly the same rate in our model, since baseline growth is deterministic. Therefore studying the price-dividend ratio allows us to remove time consideration from our analysis.

We now solve for the value of \( q_1 \) and \( q_2 \) for current period \( t \) (noting that this also gives us the price-dividend ratio for any period) by expressing all dividend and growth expressions in terms of known values at period \( t \). Note that from (3), we have

\[
G_{ij} = \delta^{j-i} G,
\]

(19)

which governs the relation between dividends of different states in neighboring time periods.
Appealing to (3) and (19), we simplify (17) to

\[
\frac{1}{\beta} = \Pi_{11} \left[ G^{-\theta} \left( \frac{q_1 G d_{1,t} + G d_{1,t}}{q_1 d_{1,t}} \right) \right] + \\
\Pi_{12} \left[ (\delta G)^{-\theta} \left( \frac{q_2 \delta G d_{1,t} + \delta G d_{1,t}}{q_1 d_{1,t}} \right) \right] \\
= \Pi_{11} \left[ G^{-\theta} \left( \frac{q_1 G + G}{q_1} \right) \right] + \Pi_{12} \left[ (\delta G)^{-\theta} \left( \frac{q_2 \delta G + \delta G}{q_1} \right) \right], \tag{20}
\]

and analogously, we simplify (18) to

\[
\frac{1}{\beta} = \Pi_{21} \left[ (\delta^{-1}G)^{-\theta} \left( \frac{q_1 \delta^{-1} G d_{2,t} + \delta^{-1} G d_{2,t}}{q_2 d_{2,t}} \right) \right] + \\
\Pi_{22} \left[ G^{-\theta} \left( \frac{q_2 G d_{2,t} + G d_{2,t}}{q_2 d_{2,t}} \right) \right] \\
= \Pi_{21} \left[ (\delta^{-1}G)^{-\theta} \left( \frac{q_1 \delta^{-1} G + \delta^{-1} G}{q_2} \right) \right] + \Pi_{22} \left[ G^{-\theta} \left( \frac{q_2 G + G}{q_2} \right) \right]. \tag{21}
\]

From (20), we have the equation

\[
q_1 \left[ 1/\beta - \Pi_{11} G^{1-\theta} \right] + q_2 \left[ -\Pi_{12} (\delta G)^{1-\theta} \right] = \Pi_{11} G^{1-\theta} + \Pi_{12} (\delta G)^{1-\theta}, \tag{22}
\]

and from (21), we have

\[
q_1 \left[ -\Pi_{21} (\delta^{-1}G)^{1-\theta} \right] + q_2 \left[ 1/\beta - \Pi_{22} G^{-\theta} \right] = \Pi_{21} (\delta^{-1}G)^{1-\theta} + \Pi_{22} G^{1-\theta}. \tag{23}
\]

We now have a system of equations for \(q_1\) and \(q_2\), which we express in the following matrix form

\[
\begin{pmatrix}
1/\beta - \Pi_{11} G^{1-\theta} & -\Pi_{12} (\delta G)^{1-\theta} \\
-\Pi_{21} (\delta^{-1}G)^{1-\theta} & 1/\beta - \Pi_{22} G^{1-\theta}
\end{pmatrix}
\begin{pmatrix}
q_1 \\
q_2
\end{pmatrix}
=
\begin{pmatrix}
\Pi_{11} G^{1-\theta} + \Pi_{12} (\delta G)^{1-\theta} \\
\Pi_{21} (\delta^{-1}G)^{1-\theta} + \Pi_{22} G^{1-\theta}
\end{pmatrix}, \tag{24}
\]

and so we can now solve for \(q_1\) and \(q_2\).
Once we have \( q_1 \) and \( q_2 \), we have \( p_{i,t} = q_i d_{i,t} \) for \( i = 1, 2 \) and \( \forall t \). Thus, to determine \( R_{i,t}^e \), we solve (13) for our two-state case, giving

\[
R_{i,t}^e = E_{i,t} \left[ \frac{p_{t+1} + d_{t+1}}{p_t} \right] \\
= \frac{1}{p_{i,t}} E_{i,t}(p_{t+1} + d_{t+1}) \\
= \frac{1}{p_{i,t}} \sum_{j=1}^{2} \Pi_{ij} (p_{j,t+1} + d_{j,t+1}) \\
= \frac{1}{q_i d_{i,t}} \sum_{j=1}^{2} \Pi_{ij} (q_j Gd_{j,t} + Gd_{j,t}), \tag{25}
\]

where our expectation operator \( E_{i,t} \) denotes expectation conditional upon all information known at period \( t \), given that we are currently in state \( i \).

For the riskless rate, from (11) we have

\[
R_{i,t}^f = \frac{d_{i,t}^{-\theta}}{\beta \sum_{j=1}^{2} \Pi_{ij} d_{j,t+1}^{-\theta}} \\
= \frac{d_{i,t}^{-\theta}}{\beta \sum_{j=1}^{2} \Pi_{ij} (Gd_{j,t})^{-\theta}}. \tag{26}
\]

Combining (25) and (26), we determine the state and period-specific equity premium as

\[
\text{Equity Premium}_{i,t} = R_{i,t}^e - R_{i,t}^f. \tag{27}
\]
4.5 Generalizing to \( n \)-states

In general, we obtain the \( n \)-state model version of (24) as

\[
\frac{1}{\beta} I_n - \begin{pmatrix}
    \Psi_{11} & \Psi_{12} & \ldots & \Psi_{1n} \\
    \Psi_{21} & \Psi_{22} & \ldots & \Psi_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    \Psi_{n1} & \Psi_{n2} & \ldots & \Psi_{nn}
\end{pmatrix}
\begin{pmatrix}
    q_1 \\
    q_2 \\
    \vdots \\
    q_n
\end{pmatrix} = \begin{pmatrix}
    \sum_{i=1}^{n} \Pi_{1i}(\delta^{i-1}G)^{1-\theta} \\
    \sum_{i=1}^{n} \Pi_{2i}(\delta^{i-2}G)^{1-\theta} \\
    \vdots \\
    \sum_{i=1}^{n} \Pi_{ni}(\delta^{i-n}G)^{1-\theta}
\end{pmatrix},
\]

(28)

where \( \Psi_{ij} = \Pi_{ij}(\delta^{j-i}G)^{1-\theta} \). Solving for \( q_i \), we can determine the price-dividend ratios in each state. In conjunction with dividend values, we can determine the equity prices in each state and period, and consequently we have for equity returns

\[
R_{i,t}^e = E_{i,t} \left[ \frac{p_{t+1} + d_{t+1}}{p_t} \right] = \frac{1}{p_{i,t}} E_{i,t}(p_{t+1} + d_{t+1})
= \frac{1}{p_{i,t}} \sum_{j=1}^{n} \Pi_{ij}(p_{j,t+1} + d_{j,t+1})
= \frac{1}{q_i d_{i,t}} \sum_{j=1}^{n} \Pi_{ij} (q_j G d_{j,t} + G d_{j,t}).
\]

(29)

Importantly, note that we can rewrite (29) in a manner that does not involve \( t \) by expressing all \( d_{j,t} \) in terms of \( d_{i,t} \). This gives us

\[
R_{i}^e = \frac{1}{q_i d_{i,t}} \sum_{j=1}^{n} \Pi_{ij} (q_j G \delta^{j-i} d_{i,t} + G \delta^{j-i} d_{i,t})
= \frac{1}{q_i} \sum_{j=1}^{n} \Pi_{ij} (q_j G \delta^{j-i} + G \delta^{j-i}),
\]

(30)
which shows us that the return on equity depends only on the state.

For the riskless return, similar to (26) we have from (11),

\[
R_{t}^f = \frac{d_{i,t}^{-\theta}}{\beta \sum_{j=1}^{n} \Pi_{ij}d_{j,t+1}^{-\theta}} \\
= \frac{d_{i,t}^{-\theta}}{\beta \sum_{j=1}^{n} \Pi_{ij}(Gd_{j,t})^{-\theta}}.
\]

We can re-express (31) in the same manner that we did with equity returns to remove all period-dependent terms, hence

\[
R_{t}^f = \frac{d_{i,t}^{-\theta}}{\beta \sum_{j=1}^{n} \Pi_{ij}(G\delta{j}^{-i}d_{i,t})^{-\theta}} \\
= \frac{1}{\beta \sum_{j=1}^{n} \Pi_{ij}(G\delta{j}^{-i})^{-\theta}},
\]

thus we see that the return on the riskless asset also depends only on the state.

Therefore, we also have a period-independent equity premium for each state from (30) and (32)

\[
\text{Equity Premium}_i = R_i^e - R_i^f.
\]

4.6 Determining Average Equity Premia

We have that (33) gives us conditional equity premia based on the current state. For any model and specification for stochastic matrix \(\Pi\), we are interested in ergodic statistics for the equity premium and returns on equity and the risk-free asset. In order to obtain such statistics, we determine the stationary probability vector \(P\) for stochastic matrix \(\Pi\), where \(P_j\)
gives the probability of being in state \( j \) in some arbitrarily long time in the future, regardless of today’s state.

As our stochastic matrix \( \Pi \) is necessarily non-negative, the Perron-Frobenius theorem guarantees that such a stationary vector \( P \) exists, and is given by

\[
\lim_{k \to \infty} (\Pi^k)_{i} = P, \quad 1 \leq i \leq n, \tag{34}
\]

where the subscript \( i \) denotes the \( i^{th} \) row of the matrix.

Note that \( P \) is independent of \( i \). Hence, we have that \( P_j \) gives the long-term probability of being in state \( j \) and we can apply it to our conditional equity premia and other statistics to determine ergodic long-term averages.

Taking the vector of equity premia conditional on states

\[
\Lambda = \begin{pmatrix}
R_{1}^{c} - R_{1}^{f} \\
R_{2}^{c} - R_{2}^{f} \\
\vdots \\
R_{n}^{c} - R_{n}^{f}
\end{pmatrix},
\tag{35}
\]

and combining it with \( P \) produces the model’s long-term average equity premium, given as

\[
\text{Equity Premium} = P \ast \Lambda. \tag{36}
\]

Along the same lines, we can determine ergodic averages for other important state-dependent variables such as asset returns and price-dividend ratio. This process gives us time-invariant statistics, which are what we consider when we generally refer to the model’s equity premium (and other statistics).
4.7 Reproducing Mehra and Prescott’s Original Results

Utilizing our established model, we can parameterize it in the manner that Mehra and Prescott (1985) calibrated their original study on the equity premium. As discussed in the literature review, Mehra and Prescott set the parameters of their model to be equivalent to those observed for the U.S. economy for 1889-1978.

We use our model to mimic the characteristics of Mehra and Prescott’s original model. To match the specifications that they used, we use a period length of one year and set $G = 1.018$, $\delta = 1.036$, and stochastic matrix

$$
\Pi = \begin{pmatrix}
0.43 & 0.57 \\
0.57 & 0.43
\end{pmatrix}.
$$

In the same manner as Mehra and Prescott, we vary the values of $\beta$ and $\theta$ within the ranges $0 < \beta < 1$ and $0 < \theta \leq 10$. We obtain near-identical results to their work, with our model calibrated to mimic theirs requiring an average return on riskless asset $R_f = 1.04$ to justify an equity premium of merely $0.32\%$.

Also in similar fashion to Mehra and Prescott, we find that implausible values for $\theta$ and $\beta$ can justify the historical equity premium but only at the consequence of inflating the returns on the risk-free asset (and eventually even those for equity) to levels far and away from what has been historically observed. Hence, we confirm that our model is robust enough to produce the same asset pricing puzzles originally described by Mehra and Prescott.

5 Implementation of Stochastic Matrices

We now apply our established model towards studying the effects of different economic disaster specifications on the magnitude of the equity premium.

We demonstrate how different specifications of $\Pi$, the stochastic probability transition
matrix, can be built to simulate variations of growth variability as well as incorporate economic disasters in our model. We explain four different stochastic matrices \( \Pi_{SS}, \Pi_G, \Pi_{MR}, \) and \( \Pi_{FD} \), each of which is more sophisticated and builds upon the design of the previous matrices.

In sections 6 and 7, we will incorporate multiple stochastic matrices together in a single model by building the possibility of multiple regimes.

5.1 Modeling Standard Variability of Consumption Growth

Each state of our model represents a different level of economic health with respect to the baseline trend. The value of \( \delta \), the ratio of dividends between neighboring states, controls the variability of consumption. Hence by specifying a single dividend value \( d_{b,t=1} \), which is the baseline state’s dividend in period \( t = 1 \), we can determine the dividend values for all other states using \( \delta \). Furthermore, since all dividend values increase by growth multiple \( G \) after each successive period, we can also determine all future dividend values for all states from \( d_{b,t=1} \).

In order to be able to simulate both normal variation in asset returns as well as potential severe falls caused by a depression or disaster, we employ a version of the model with a considerably larger number of states. Using \( n > 100 \) states allows us to properly simulate both Brownian motion in asset returns as well as extreme, single-period crashes in dividend levels. Here, we detail possible constructions for the stochastic matrix that yield more realistic simulations of asset returns in reality.

5.1.1 The Simple Symmetric Random Walk

A basic model for the normal variation in consumption returns uses a probability transition matrix \( \Pi_{SS} \) which simulates a simple symmetric random walk movement through different states over time. In combination with the baseline growth created by \( G \), the normal state
transitions comprise a random walk with drift. Such a random walk probability transition matrix takes the form

$$
\Pi_{SS} = \begin{pmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 & \ldots & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & \frac{1}{2} & \ldots & 0 & 0 & 0 & 0 \\
: & : & : & : & \ldots & : & : & : & : \\
0 & 0 & 0 & 0 & \ldots & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
0 & 0 & 0 & 0 & \ldots & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 & \ldots & 0 & 0 & \frac{1}{2} & \frac{1}{2}
\end{pmatrix}.
$$

With our simple symmetric random walk process, in every state $i$, there is a $1/2$ chance of being in either state $i - 1$ or state $i + 1$ in the next period. The exception is in states 1 and $n$, where the random walk process is barriered at the first and last state, due to the finite nature of our discrete state model.

Such a stochastic matrix treatment models the simplest possible random walk: the next period’s state is either one above or one below the current period’s state, each with probability $1/2$. Consequentially, the next period’s dividend is either one additional standard deviation above or below the current period’s dividend relative to the baseline value (hence the value of $\delta$ here controls the standard deviation of growth). Of course, variations of actual asset returns on a year-to-year basis are by no means bound to an upper limit of a single standard deviation, so further refinement of our stochastic matrix is warranted.

5.1.2 The Gaussian Random Walk

In order to consider a more sophisticated treatment of Brownian motion where larger jumps are possible, consider a state transition where instead of just one state up or down, there exists the potential to move up to three states up or down (representing one, two, and
This describes an example of a Gaussian random walk, a common basic model for financial market data where the frequency of different step sizes varies according to a normal distribution.

Here, we have our stochastic matrix modeling such a Gaussian random walk driving the state transition, with jumps of up to three states possible.

\[
\Pi_G = \begin{pmatrix}
\frac{1}{2} & s_1 & s_2 & s_3 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & s_1 & s_2 & s_3 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
s_2 + \frac{s_3}{2} & s_1 + \frac{s_3}{2} & 0 & s_1 & s_2 & s_3 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
s_3 & s_2 & s_1 & 0 & s_1 & s_2 & s_3 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & s_3 & s_2 & s_1 & 0 & s_1 & s_2 & s_3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & s_3 & s_2 & s_1 & 0 & s_1 + \frac{s_3}{2} & s_2 + \frac{s_3}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & s_3 & s_2 & s_1 & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & s_3 & s_2 & s_1 & \frac{1}{2}
\end{pmatrix}
\]

Note that the minimum dimension for such a matrix is \((7 \times 7)\). Here, the values for \(s_1\), \(s_2\), and \(s_3\) represent the probabilities for one, two, and three-standard deviation events occurring between periods. As we maintain that such movements are normally distributed, \(s_1 = 0.34\), \(s_2 = 0.135\), and \(s_3 = 0.025\).

Note that in comparison to the simple symmetric random walk stochastic matrix \(\Pi_{SS}\), where the absolute change in the dividend is always by a factor of \(\delta\) (hence \(\delta - 1\) is equal to the standard deviation of growth), here we have the lower value \(\delta = 1.0174\) in order to match the historical standard deviation of growth. See section C in the appendix for the calculation of \(\delta\) for \(\Pi_G\).

Due to the barriered nature of the model’s top-most and bottom-most states, returns on equity and on the riskless asset differ abnormally less in those states. This is because
the uncertainty in the top-most and bottom-most states is much lower compared to all other states, where both upwards and downwards movement is possible. Having a low number of total states increases the probability that the top-most or bottom-most states will be encountered in the course of stochastic transformations, hence we want to use \( n \) high enough such that the model does not become overly influenced by the behavior of the barriered states, but also not so high that we introduce more computational intensiveness than necessary. Table 1 shows the model results for a variety of possible \( n \) number of states, varying from \( n = 25 \) to \( n = 501 \).

Given our results in Table 1, we see that as we increase the number of states in our model, the long-run average return on equity and the riskless asset both decrease, with the return on equity decreasing at a slightly faster rate. This results in a lower equity premium for \( n = 501 \), compared to \( n = 25 \), although the difference is a marginal 8 basis points. The long-run average price-dividend ratio increases slowly with more states, although the decrease becomes vanishingly small with larger \( n \). As a final note, if we increase \( n \) past 1000, the results are negligibly different than for \( n = 501 \).

Because our model considers one year to be the simulated time length of one period (and hence one stochastic transition), and we look to simulate the U.S. economy over the the 84-year period of 1926-2010, having potentially more than 75 states above and below the mean trend does not further enrich our model because we would not expect a movement of 75 standard deviations (or even anything near that size of deviation) above or below the mean trend over the course of 84 simulated period transitions, even in a random walk. Hence, such intuition in combination with our results in Table 1 which show that increasing \( n \) quickly comes to have inconsequential impact on results leads us to designate \( n = 151 \) as a maximum number of states to use.
5.1.3 Building Mean Reversion

The third characteristic we look to add to our stochastic probability transitions is mean reversion. With our previous two treatments, our model’s probability transitions are state-invariant; they do not depend on what state (or relative state) the economy is currently in. Of course in reality we would expect, for developed nations, an economy well above trend to be more likely to see slower growth in the near future compared to an economy near the baseline level and especially compared to an economy well below trend. However, our $\Pi_{SS}$ and $\Pi_{G}$ as detailed previously do not exhibit such mean reverting behavior, being equally likely to go up or down regardless of where relative to the baseline level the current state is.

Such considerations motivate us to create stochastic transition matrix $\Pi_{MR}$ which tends to keep the model close to the baseline dividend level. To do so, we modify $\Pi_{G}$ with a mechanism that skews the probabilities in favor of moving to higher states when below the mean, and in favor of moving to lower states when above the mean.

We denote the baseline state to be the middle of the possible state transitions, hence $i_{b} = \lceil n/2 \rceil$. For symmetry purposes, we necessarily use an odd value for $n$ so that there may be an equal number of potential states above and below the middle, baseline state. Note that this baseline state $i_{b}$ is also the state that receives the initial baseline dividend value $d_{b,t=1}$.

In order to build mean reversion into $\Pi_{G}$, we modify the probabilities depending on their distance from the middle of the stochastic matrix. We multiply each row’s downward transition probabilities by a scaling factor

$$\omega_{D,i} = 1 + \frac{i - i_{b}}{\lceil n/2 \rceil}, \quad (38)$$
and multiply each row’s upward transition probabilities by an offsetting scaling factor

\[
\omega_{U,i} = 1 + \frac{i_b - i}{\left\lceil n/2 \right\rceil}.
\]  (39)

Note that \(\omega_D + \omega_U = 2\) and \(\omega_D, \omega_U \in [0, 2]\) for \(i \in \{1, 2, \ldots n\}\). Hence for any given state \(i\), the transition probabilities for the current period to the next are given by row \(i\) in the stochastic matrix and are

\[
\Pi_{MR,i} = \left( \ldots 0 \ \omega_D s_3 \ \omega_D s_2 \ \omega_D s_1 \ 0 \ \omega_U s_1 \ \omega_U s_2 \ \omega_U s_3 \ 0 \ \ldots \right). \]  (40)

Note that the extreme upper-left and lower-right hand corners of our matrix are special cases and we must treat them uniquely in order to maintain the properties of the stochastic matrix. See section B.3 in the appendix for an explicit formulation of the special cases for \(\Pi_{MR}\).

Our scaling factor linearly decreases the probabilities of moving away from the mean as we move farther in that direction while increasing the probabilities of moving towards the mean by a corresponding amount. Hence at lowest state \(i = 1\), we have \(\omega_D = 0\) and \(\omega_U = 2\); consequently the probability moving to a higher state is 1 while the probability of moving to a lower state is 0, whereas at highest state \(i = n\), the opposite is true. In a model with 15 states, the eighth state acts as the baseline state and the fifth state, for example, has its downwards transition probabilities scaled by a factor of \(4/7\) and its upwards transition probabilities scaled by a factor of \(10/7\). This encourages return towards the mean state, and such mean reverting behavior increases in strength the farther we are from the mean state in the matrix.
5.2 Modeling Rare Economic Disasters

Rietz (1988) first proposes a scenario when, in any given period, there exists a low probability of consumption severely contracting in the next period. Such a contraction simulates the occurrence of an economic depression and the corresponding plunge in the value of equity and real GDP per capita. Barro (2006) further extends Rietz’s idea of the low-probability economic crash and empirically analyzes the frequency and magnitude of the historical depressions that occurred in developed nations in the 20th century. Here, we develop a similar idea which allows our model to have rare, large falls in dividend levels.

In our multi-state discrete model, we represent depressions through having a low probability of a multi-state fall in the state transition between periods. In order to generate a contraction of size \( b \), we require a single-period fall of \( z \) states, where

\[
z = \lceil -\log_\delta(1 - b) \rceil. \tag{41}
\]

Hence, in order to model a contraction of size \( b \), our model must, as a minimum, have enough states \( n \) such that \( n > z \). Furthermore, due to our stochastic matrix’s mean reversion tends to keep the state close to \( i_b \) in the absence of shocks, it is beneficial for our model to have \( i_b > z \), which is approximately equivalent to \( n > 2z \).

In order to generate the most severe historical disasters empirically examined by Barro (2006) where \( b = 0.64 \), we have \( z = 36 \).\(^9\) Hence, \( n = 151 \) would generally be a sufficient size for our stochastic matrix, unless we wanted to study disasters of exceptionally extreme magnitude where we fall by a hundred states or more.

Rietz (1988) and Barro (2006) model disasters using two variables: single-period disaster

---

\(^8\)We follow Barro’s notation for contraction size, where contraction fraction \( 0 < b < 1 \) represents the current period’s dividend being a factor of \( (1 - b) \) times the previous period’s in the event of a disaster. For example, \( b = 0.25 \) would represent a single-period dividend fall of 25% when a disaster occurs.

\(^9\)This uses \( \delta = 1.0174 \). This is because the average absolute change between states is equivalent to the Gaussian random walk model.
probability \( p \) and contraction fraction \( b \). These two variables control the frequency and magnitude of the disasters that occur within their models. For both Rietz and Barro, \( p \) is a constant. Rietz treats \( b \) as a constant as well, while Barro allows \( b \) to have a distribution based on historical data (although the focus is on the expectation of \( b \)).

We can modify our mean-reverting Gaussian random walk stochastic matrix \( \Pi_{MR} \) to incorporate the possibility of low-probability disasters. After specifying a disaster probability \( p \) and disaster size \( b \) (which also defines \( z \)), our algorithm to build disasters into \( \Pi_{MR} \) operates as follows:

1. Multiply the probabilities in each row of \( \Pi_{MR} \) by \( 1 - p \). This scales down the probabilities of the normal random walk transitions in order to allow for low-probability disaster transitions.

2. For each row \( i \) in \( \Pi_{MR} \), increase the value of \( \Pi_{i,i-z} \) by \( p \) for \( i - z \geq 1 \), hence adding a probability \( p \) of falling \( z \) states given that we are currently in state \( i \) for all states, which creates our disaster mechanism. If \( i - z < 1 \), then instead increase \( \Pi_{i1} \) by \( p \), which represents a fall into the lowest possible state in our stochastic matrix.

We refer to our resulting stochastic matrix as \( \Pi_{FD} \), as it models disasters that are fixed in both scope and rarity.

In Table 2 of the appendix, we present the model’s results for using \( \Pi_{FD} \). We designate a set of baseline parameters\(^{10}\) and then vary them in the same manner as Barro (2006), studying their effects on returns and the equity premium. We see that a much more substantial equity premium appears when the model accounts for disasters, and both returns on equity and the riskless asset are lower. This yields Barro’s results as a special case of our model.

Our baseline equity premium of 2.7% is much more reasonable than the equity premium of 0.4% when not incorporating disasters. Furthermore, both the high \( p \) and high \( b \) cases

\(^{10}\)The baseline values used for \( p \) and \( b \) come from Barro’s analysis on the size and frequency of 20\textsuperscript{th} century economic depressions in OECD nations.
generate an equity premium in the neighborhood of 4%. Importantly, a 4% equity premium in the absence of firm leverage matches historical results. Barro explains how the introduction of a historically observed degree of firm leverage increases the equity premium by a factor of 1.5, hence an unlevered equity premium of 4% equates to a levered equity premium of 6%.

The primary disconnect that remains with $\Pi_{FD}$ is that the model’s return on the riskless asset is still too high compared to historical data. We continue to refine our model.

6 Incorporating Stochastic Regimes

Our previous general approach to the modeling of economic cycles considers a model where each state represents some level of the economy, with the middle state being the average baseline level. Our stochastic matrix $\Pi$ controls how our model moves between different states.

Now, we extend our model so that there can potentially be $k$ different stochastic matrices, each of which represents a different set of state transition probabilities. We refer to each of these different possible matrices as a regime. From an economic perspective, an example two-regime model could simulate a possible economy where the healthy regime operates under $\Pi_{MR}$ while the weakened regime operates under $\Pi_{FD}$.

This adds a $(k \times k)$ regime stochastic matrix $\Phi$ such that $\Phi_{xy}$ denotes the probability of being in regime $y$ in the next period, given that the current period’s regime is regime $x$. After each period, we first stochastically determine what the next period’s regime $y$ will be given the current regime $x$ using $\Phi_{xy}$, and then use the corresponding state stochastic matrix $\Pi_x$ to determine the next period’s state given the current period’s state.

Here, we explain how we extend our model to incorporate multiple regimes of economic health, each with its own different transition probabilities between states. In particular, we will see that the $n$-state model that we build in section 4.5 becomes the single-regime case.
of our general forthcoming k-regime model.

6.1 Equity Returns Under Multiple Regimes

With \( k \) possible regimes, our first-order condition (6) for equity prices in equilibrium becomes

\[
p_{i,x,t} u'(d_{i,t}) = \beta \sum_{y=1}^{k} \Phi_{xy} \sum_{j=1}^{n} \Pi_{y,ij} [u'(d_{j,t+1})(p_{j,y,t+1} + d_{j,t+1})],
\]

where \( p_{i,x,t} \) is the price of the equity share in state \( i \), regime \( x \), and period \( t \), \( \Pi_{y,ij} \) is the probability of being in state \( j \) next period given that we are in state \( i \) in the current period and regime next period \( y \), and \( \Phi_{xy} \) is the probability that the next period is under regime \( y \) given that the current one is under regime \( x \).

Applying our constant relative risk aversion utility function (7) and multiplying both sides by \( d_{i,t}^{\theta}/(\beta p_{i,x,t}) \), we have

\[
\frac{1}{\beta} = \sum_{y=1}^{k} \Phi_{xy} \sum_{j=1}^{n} \Pi_{y,ij} \left[ G_{ij}^{\theta} \left( \frac{p_{j,y,t+1} + d_{j,t+1}}{p_{j,y,t+1} + d_{j,t+1}} \right) \right],
\]

where \( G_{ij} = d_{j,t+1}/d_{i,t} \).

Equation (43) is the generalized \( k \)-regime version of (9). Hence we take a similar approach to the one we used in section 4 to solve for the price-dividend ratios. Because we now have \( k \)-regimes along with \( n \)-states, there are \( n \times k \) total price-dividend ratios to solve for, where \( q_{i,x} \) is the price-dividend ratio in regime \( x \) and state \( i \). Analogously to the single-regime model, we have time-invariant price-dividend ratios unique for each state. With the addition of regimes, each regime now has its own set of state-specific price-dividend ratios. Each state has a time-invariant dividend process for each regime because we now have \( k \) regime-specific
stochastic matrices.

The $k$-regime, $n$-state model version of (28) is

$$
\frac{1}{\beta} I_{n+k} - \begin{pmatrix}
\Psi_{11} & \Psi_{12} & \ldots & \Psi_{1,n+k} \\
\Psi_{21} & \Psi_{22} & \ldots & \Psi_{2,n+k} \\
\vdots & \vdots & \ddots & \vdots \\
\Psi_{n+k,1} & \Psi_{n+k,2} & \ldots & \Psi_{n+k,n+k}
\end{pmatrix} \begin{pmatrix}
q_{1,1} \\
q_{1,2} \\
\vdots \\
q_{1,k} \\
q_{2,1} \\
\vdots \\
q_{n,k}
\end{pmatrix},
$$

where $\Psi_{ab} = \Phi_{xy} \Pi_{y,ij} (\delta^{j-i} G)^{1-\theta}$, and $x = k - (a \mod k)$, $y = k - (b \mod k)$, $i = \lceil a/k \rceil$, $j = \lceil b/k \rceil$.

The structure of $\Psi$ superficially appears very complex, but there is a systematic pattern behind it. Here, we instructively give the explicit formulation of $\Psi$ for the $k = 2$ regimes,
\( n = 2 \) states case:

\[
\Psi = \begin{pmatrix}
\Phi_{11} \Pi_{1,11} G^{1-\theta} & \Phi_{12} \Pi_{2,11} G^{1-\theta} & \Phi_{11} \Pi_{1,12} (\delta G)^{1-\theta} & \Phi_{12} \Pi_{2,12} (\delta G)^{1-\theta} \\
\Phi_{21} \Pi_{1,11} G^{1-\theta} & \Phi_{22} \Pi_{2,11} G^{1-\theta} & \Phi_{21} \Pi_{1,12} (\delta G)^{1-\theta} & \Phi_{22} \Pi_{2,12} (\delta G)^{1-\theta} \\
\Phi_{11} \Pi_{1,21} (\delta^{-1} G^{1-\theta}) & \Phi_{12} \Pi_{2,21} (\delta^{-1} G^{1-\theta}) & \Phi_{11} \Pi_{1,22} G^{1-\theta} & \Phi_{12} \Pi_{2,22} G^{1-\theta} \\
\Phi_{21} \Pi_{1,21} (\delta^{-1} G^{1-\theta}) & \Phi_{22} \Pi_{2,21} (\delta^{-1} G^{1-\theta}) & \Phi_{21} \Pi_{1,22} G^{1-\theta} & \Phi_{22} \Pi_{2,22} G^{1-\theta}
\end{pmatrix}.
\]

Note that the structure of \( \Psi \) is almost identical to the Kronecker product of \((k \times k)\) matrix \( \Phi \) and the regime-free \( n \)-state version of \((n \times n)\) matrix \( \Psi \) from (28), except that because we now have multiple state stochastic matrices \( \Pi \), the particular stochastic matrix \( \Pi_{y} \) used in each entry depends on the next regime \( y \) of the regime stochastic matrix entry \( \Phi_{xy} \). The method of calculating any specific entry \( \Psi_{ab} \) given above allows us to derive the proper regime transition probability \( \Phi_{xy} \), state stochastic matrix \( \Pi_{y} \), and state transition probability \( \Pi_{y,ij} \) from only the values of \( a \) and \( b \).

Solving for our \( q_{i,x} \), we have the price-dividend ratios in each state under each regime. In conjunction with dividend values, we can determine the equity prices in each state and regime. Generalizing from (29), we consequently have for equity returns

\[
R_{i,x,t}^e = E_{i,x,t} \left[ \frac{p_{t+1} + d_{t+1}}{p_t} \right] = \frac{1}{p_{i,x,t}} E_{i,x,t} (p_{t+1} + d_{t+1}) = \frac{1}{p_{i,x,t}} \sum_{y=1}^{k} \Phi_{xy} \sum_{j=1}^{n} \Pi_{y,ij} (p_{j,y,t+1} + d_{j,t+1}) = \frac{1}{q_{i,x} d_{i,t}} \sum_{y=1}^{k} \Phi_{xy} \sum_{j=1}^{n} \Pi_{y,ij} (q_{j,y} G d_{j,t} + G d_{j,t}),
\] (45)
with period-independent expression

\[ R_{i,x} = \frac{1}{q_{i,x}} \sum_{y=1}^{k} \Phi_{xy} \sum_{j=1}^{n} \Pi_{y,ij} (q_{j,y} G\delta^{j-i} + G\delta^{j-i}). \]  

(46)

6.2 Riskless Asset Returns Under Multiple Regimes

For the riskless return, generalizing from the regime-less \( n \)-state expression (31) we have

\[ R_{f_{i,x,t}} = \frac{d_{i,t}^{-\theta}}{\beta \sum_{y=1}^{k} \Phi_{xy} \sum_{j=1}^{n} \Pi_{y,ij} d_{j,t+1}^{-\theta}}. \]

\[ = \frac{d_{i,t}^{-\theta}}{\beta \sum_{y=1}^{k} \Phi_{xy} \sum_{j=1}^{n} \Pi_{y,ij} (Gd_{j,t})^{-\theta}}, \]  

with period-independent expression

\[ R_{e_{i,x}} = \frac{1}{\beta \sum_{y=1}^{k} \Phi_{xy} \sum_{j=1}^{n} \Pi_{y,ij} (G\delta^{j-i})^{-\theta}}. \]  

(48)

6.3 Average Equity Premia Under Multiple Regimes

With (46) and (48), we now have an expression for the equity premium

\[ \text{Equity Premium}_{i,x} = R_{e_{i,x}} - R_{f_{i,x}}, \]  

(49)

where the premium depends on the state \( i \), regime \( x \), but not on the period.

Using the same techniques from section 4.6, we determine stationary vectors \( P_x \) for each regime \( x \)'s stochastic matrix \( \Pi_x \) as well as the stationary vector \( P_\Phi \) for the regime transition
matrix. Similar to (34), we have

$$\lim_{k \to \infty} (\Phi^k)_i = P_\Phi, \quad 1 \leq i \leq k. \quad (50)$$

Then the long-term average equity premium for our $n$-state, $k$-regime model is

$$\text{Equity Premium} = \sum_{x=1}^{n} P_{\Phi,x} \sum_{i=1}^{n} P_{x,i}(R_{i,x}^e - R_{i,x}^f). \quad (51)$$

Long-term average returns on equity and the riskless asset are determined similarly.

7 Conditional Stochastic Disasters

With regimes, we can now build economic models of an additional order of complexity compared to when we had only state transitions.

For example, consider the scenario of $k = 2$ regimes where the first regime operates under $\Pi_{\text{MR}}$ while the second regime is a guaranteed fall of $z$ states as given by (41), depending on the value of contraction fraction $b$. Hence, $\Pi_{2,ij} = 1$ for $j = \text{Max}(i - z, 1)$. We have regime stochastic matrix

$$\Phi = \begin{pmatrix} 1 - p & p \\ 1 - p & p \end{pmatrix},$$

for some disaster probability $p$.

Note we have defined a scenario identical to $\Pi_{\text{FD}}$, where there is a $p$ probability in each period of a disaster and dividends falling by fraction $b$.

Without regimes, we are forced to work within one stochastic matrix and require complicated algorithms such as the one we used in section 5.2 to create $\Pi_{\text{FD}}$. But with regimes, we can much more easily construct such rare disaster scenarios. Furthermore, we now have the mathematical machinery capable of building stochastic mechanisms that are not possible
without multiple regimes.

7.1 A Conditional Multi-Regime Model

Consider our discussion in section 2, where we noted the possibility of recessionary after-shocks following a serious economic depression. In fact, it is natural to think that the occurrence of a depression weakens the economy, which raises the probability of $p$ while decreasing the value of $b$. After some time, the economy recovers and the values of $p$ and $b$ return to what they are in a health economy.

It is very difficult to simulate such a stochastic model using a single probability transition matrix, but with the application of multiple regimes it becomes relatively simple.

Let $k = 4$ regimes. Regimes 1 and 3 represent a healthy and weakened economy, respectively. Regime 2 is the depression regime while regime 4 is a recession regime, reachable as an aftershock from the weakened economy. Regimes 1 and 3 operate under $\Pi_{MR}$. Regime 2 is a guaranteed fall of $z_1$ states as given by (41), depending on the value of disaster contraction fraction $b_1$. Regime 4 is a guaranteed fall of $z_2$ states, controlled by a different contraction fraction $b_2$.

We define two separate disaster probabilities $p_d$ and $p_r$, where $p_d$ is the probability of a depression occurring and $p_r$ is the probability of a recession. Our regime stochastic matrix is

$$
\Phi = \begin{pmatrix}
1 - p_d & p_d & 0 & 0 \\
0 & 0 & 1 - p_r & p_r \\
r & 0 & 1 - r - p_r & p_r \\
0 & 0 & 1 - p_r & p_r
\end{pmatrix}.
$$

Hence, there is a $p_d$ probability of a depression occurring during a healthy economy. If a depression occurs, it throws the economy into a weakened regime, where there is a $p_r$ probability of a recession occurring in a given period. While the economy is weakened, there
is an $r$ probability of the economy recovering in any given period, at which point our model returns to the healthy regime. The average length of an economic recovery defines the value of $r$ (our baseline parameter of $r = 0.2$ represents an average length of full economic recovery from a depression of five years).

In order to consider the implications of our multi-regime model, we compare its results to those observed for the fixed disaster model. While the fixed disaster model has an average long-term real return of 7.3% for baseline parameters, our multi-regime model produces a much more reasonable long-term real return of 5% for baseline parameters, with only a 0.27% decrease in the equity premium.

The high $b_d$ case generates the historical unlevered equity premium of 4% with a return on the riskless asset of 2.55%. Compare these results to the analogous ‘high $b$’ case for $\Pi_{FD}$, which produced a similar equity premium of 4.4% but an implausibly large riskless return of 5.3%.

In summary, our conditional multi-regime model is able to produce the historical equity premium of 6% (once firm leverage is applied). It does so while maintaining a plausibly low return on the riskless asset.

8 Concluding Remarks

The results of our multi-regime model are encouraging, taking another step towards resolving the equity premium puzzle within its original economic framework. We achieve a considerably more reasonable real return compared to the fixed disasters model while negligibly affecting the model’s equity premium.

More importantly, our generalized $k$-regime, $n$-state model allows the study of significantly more sophisticated stochastic processes. Shocks to growth no longer need to be independent and identically distributed as in Barro’s study; we now have closed-form, com-
putable solutions for a substantially larger class of possible economic cycles.

This paper focuses on creating the mathematical tools and demonstrating their application, flexibility, and robustness. An extension would be to conduct a statistical study of depressions and recessions in order to build and parameterize multi-regime models which closely approximate observed historical market structure.
9 References


A Tables of Results

Table 1: Results for $\Pi_G$ when varying number of states $n$.

<table>
<thead>
<tr>
<th>$n$ (number of states)</th>
<th>25</th>
<th>51</th>
<th>101</th>
<th>151</th>
<th>251</th>
<th>501</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$ (ratio of dividend between states)</td>
<td>1.0174</td>
<td>1.0174</td>
<td>1.0174</td>
<td>1.0174</td>
<td>1.0174</td>
<td>1.0174</td>
</tr>
<tr>
<td>$\theta$ (coeff. of relative risk aversion)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$G$ (growth rate)</td>
<td>1.0205</td>
<td>1.0205</td>
<td>1.0205</td>
<td>1.0205</td>
<td>1.0205</td>
<td>1.0205</td>
</tr>
<tr>
<td>$\beta$ (time discount factor)</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>$R_e$ (long-term average equity return)</td>
<td>1.1060</td>
<td>1.1049</td>
<td>1.1043</td>
<td>1.1042</td>
<td>1.1041</td>
<td>1.1040</td>
</tr>
<tr>
<td>$R_f$ (long-term average riskless return)</td>
<td>1.1016</td>
<td>1.1012</td>
<td>1.1011</td>
<td>1.1010</td>
<td>1.1010</td>
<td>1.1009</td>
</tr>
<tr>
<td>long-term average equity premium</td>
<td>0.0044</td>
<td>0.0037</td>
<td>0.0032</td>
<td>0.0032</td>
<td>0.0031</td>
<td>0.0031</td>
</tr>
</tbody>
</table>

Table 2: Results for $\Pi_{FD}$ when varying differing parameters.

<table>
<thead>
<tr>
<th>$n$ (number of states)</th>
<th>baseline</th>
<th>low $\theta$</th>
<th>high $p$</th>
<th>high $b$</th>
<th>no disasters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$ (ratio of dividend between states)</td>
<td>1.0174</td>
<td>1.0174</td>
<td>1.0174</td>
<td>1.0174</td>
<td>1.0174</td>
</tr>
<tr>
<td>$\theta$ (coeff. of relative risk aversion)</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$G$ (growth rate)</td>
<td>1.0205</td>
<td>1.0205</td>
<td>1.0205</td>
<td>1.0205</td>
<td>1.0205</td>
</tr>
<tr>
<td>$\beta$ (time discount factor)</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>$b$ (contraction fraction)</td>
<td>0.29</td>
<td>0.29</td>
<td>0.29</td>
<td>0.35</td>
<td>-</td>
</tr>
<tr>
<td>$p$ (disaster probability)</td>
<td>0.017</td>
<td>0.017</td>
<td>0.025</td>
<td>0.017</td>
<td>0</td>
</tr>
<tr>
<td>$R_e$ (long-term average equity return)</td>
<td>1.0995</td>
<td>1.0834</td>
<td>1.0987</td>
<td>1.0974</td>
<td>1.1054</td>
</tr>
<tr>
<td>$R_f$ (long-term average riskless return)</td>
<td>1.0729</td>
<td>1.0676</td>
<td>1.0617</td>
<td>1.0531</td>
<td>1.1010</td>
</tr>
<tr>
<td>long-term average equity premium</td>
<td>0.0266</td>
<td>0.0158</td>
<td>0.0370</td>
<td>0.0443</td>
<td>0.0044</td>
</tr>
</tbody>
</table>
Table 3: Results for a multi-regime model with regimes of healthy economy, weakened economy, depression, and recession.

<table>
<thead>
<tr>
<th>Multi-Regime Conditional Disasters Model Results</th>
<th>baseline</th>
<th>high $p_d$</th>
<th>high $b_d$</th>
<th>low $b_r$</th>
<th>low $r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$ (number of states)</td>
<td>151</td>
<td>151</td>
<td>151</td>
<td>151</td>
<td>151</td>
</tr>
<tr>
<td>$\delta$ (ratio of dividend between states)</td>
<td>1.0174</td>
<td>1.0174</td>
<td>1.0174</td>
<td>1.0174</td>
<td>1.0174</td>
</tr>
<tr>
<td>$\theta$ (coeff. of relative risk aversion)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$G$ (growth rate)</td>
<td>1.0205</td>
<td>1.0205</td>
<td>1.0205</td>
<td>1.0205</td>
<td>1.0205</td>
</tr>
<tr>
<td>$\beta$ (time discount factor)</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>$b_d$ (depression contraction fraction)</td>
<td>0.29</td>
<td>0.29</td>
<td>0.35</td>
<td>0.29</td>
<td>0.29</td>
</tr>
<tr>
<td>$p_d$ (depression probability)</td>
<td>0.017</td>
<td>0.025</td>
<td>0.017</td>
<td>0.017</td>
<td>0.017</td>
</tr>
<tr>
<td>$b_r$ (recession contraction fraction)</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.145</td>
<td>0.2</td>
</tr>
<tr>
<td>$p_r$ (recession probability)</td>
<td>0.051</td>
<td>0.051</td>
<td>0.051</td>
<td>0.051</td>
<td>0.051</td>
</tr>
<tr>
<td>$r$ (recovery probability)</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$R^e_{av}$ (long-term average equity return)</td>
<td>1.0738</td>
<td>1.0639</td>
<td>1.0661</td>
<td>1.0759</td>
<td>1.0722</td>
</tr>
<tr>
<td>$R^f_{av}$ (long-term average riskless return)</td>
<td>1.0499</td>
<td>1.0306</td>
<td>1.0255</td>
<td>1.0516</td>
<td>1.0488</td>
</tr>
<tr>
<td>long-term average equity premium</td>
<td>0.0239</td>
<td>0.0333</td>
<td>0.0406</td>
<td>0.0243</td>
<td>0.0234</td>
</tr>
<tr>
<td>long-term average price-dividend ratio</td>
<td>17.464</td>
<td>20.062</td>
<td>19.709</td>
<td>16.967</td>
<td>17.816</td>
</tr>
</tbody>
</table>
B Stochastic Matrices

B.1 $\Pi_{SS}$

$$\Pi_{SS} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \ldots & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \ldots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$  

B.2 $\Pi_G$

$$\Pi_G = \begin{pmatrix} \frac{1}{2} & s_1 & s_2 & s_3 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & s_1 & s_2 & s_3 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 \\ s_2 + \frac{s_3}{2} & s_1 + \frac{s_3}{2} & 0 & s_1 & s_2 & s_3 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 \\ s_3 & s_2 & s_1 & 0 & s_1 & s_2 & s_3 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & s_3 & s_2 & s_1 & 0 & s_1 & s_2 & s_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & s_3 & s_2 & s_1 & 0 & s_1 + \frac{s_3}{2} & s_2 + \frac{s_3}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & s_3 & s_2 & s_1 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & s_3 & s_2 & s_1 & \frac{1}{2} \end{pmatrix}.$$
### B.3  $\Pi_{MR}$

Rows 1 through 4:

$$
\Pi_{MR} = \begin{pmatrix}
0 & \omega_{U,1}s_1 & \omega_{U,1}s_2 & \omega_{U,1}s_3 & 0 & \ldots \\
(s_1 + s_2 + s_3)\omega_{D,2} & 0 & \omega_{U,2}s_1 & \omega_{U,2}s_2 & \omega_{U,2}s_3 & \ldots \\
\omega_{D,3} \left( s_2 + \frac{s_3}{2} \right) & \omega_{D,3} \left( s_1 + \frac{s_3}{2} \right) & 0 & \omega_{U,3}s_1 & \omega_{U,3}s_2 & \ldots \\
\omega_{D,4}s_1 & \omega_{D,4}s_2 & \omega_{D,4}s_3 & 0 & \omega_{U,4}s_1 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
$$

Rows $n - 3$ through $n$:

$$
\Pi_{MR} = \begin{pmatrix}
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix}
$$

$$
\begin{pmatrix}
\ldots \omega_{D,n-3}s_1 & 0 & \omega_{U,n-3}s_1 & \omega_{U,n-3}s_2 & \omega_{U,n-3}s_3 \\
\ldots \omega_{D,n-2}s_2 & \omega_{D,n-2}s_1 & 0 & \omega_{U,n-2} \left( s_1 + \frac{s_3}{2} \right) & \omega_{U,n-2} \left( s_2 + \frac{s_3}{2} \right) \\
\ldots \omega_{D,n-1}s_3 & \omega_{D,n-1}s_2 & \omega_{D,n-1}s_1 & 0 & \left( s_1 + s_2 + s_3 \right)\omega_{U,n-1} \\
\ldots 0 & \omega_{D,n}s_3 & \omega_{D,n}s_2 & \omega_{D,n}s_1 & 0
\end{pmatrix}
$$

Note that $\omega_{D,1} = \omega_{U,n} = 0$. 
C Determination of Model Variables

In order to determine values for our model’s \( g \) and \( \delta \), we look to historical data. During the 84-year period of 1926-2009, real GDP per capita in the U.S. increased at an annualized rate of 2.05\%,\(^{11}\) with a standard deviation of \( \sigma = 2.4\% \). This gives us \( G = 1.0205 \).

As we use a Gaussian random walk as the basic probability transition specification, our value of \( \delta \) varies based on the historical standard deviation but is not simply equal to \( \sigma \), due to larger transitions than a single standard deviation being possible. Note that in our mean reverting specification (40), \( \omega_Ds_k + \omega_Ds_k = 2s_k \) for \( k = 1, 2, 3 \), so the average transition movement size is not affected because the total weight of each magnitude of transition remains unchanged.

In order to determine the value of \( \delta \) so that the average single-period variation of dividends is equal to the historical standard deviation \( \sigma \), we solve

\[
\frac{s_1 \delta + s_2 \delta^2 + s_3 \delta^3}{s_1 + s_2 + s_3} - 1 = \sigma, \tag{52}
\]

and note that we specified \( s_1 = 0.34, s_2 = 0.135, \) and \( s_3 = 0.025 \) so that our stochastic matrix models a Gaussian random walk. Solving (52) for \( \delta \) using \( \sigma = 0.024 \), we have \( \delta = 1.0174 \), which we use for all of our models based on a three-step Gaussian random walk.