Is The United States A Lucky Survivor: A Hierarchical Bayesian Approach*

Jules van Binsbergen[†] Sophia Hua[‡] Jonas Peeters[§]

Jessica A. Wachter[¶]

May 2024

ABSTRACT

Using international data, we quantify the magnitude of survivorship bias in U.S. equity market performance, and find that it explains about one third of the equity risk premium in the past century. We model the subjective crash belief of an investor who infers the crash risk in the U.S. by cross learning from other countries. The U.S. crash probability shows a persistent and widening divergence from the implied global average. We attribute the upward bias in the measured equity premium to crashes that did not occur in-sample and to positive shocks to valuations resulting from learning about the probability.

*We thank Sid Chib, Ian Dew-Becker, Winston Dou, Alice Gindin, Moritz Lenel, Karen Lewis,Nick Roussanov, Frank Schorfheide, Rob caert Stambaugh, and seminar participants at Wharton, NBER-NSF SBIES, the Econometric Society Winter Meeting, the 5-star Conference, and the WFA for comments. The authors have nothing to disclose

[†]The Wharton School, University of Pennsylvania and NBER; julesv@wharton.upenn.edu.

[‡]BlackRock;

[§]The Wharton School, University of Pennsylvania

[¶]The Wharton School, University of Pennsylvania and the Securities and Exchange Commission; jwachter@wharton.upenn.edu. This paper was initially released prior to the author joining the Commission. The Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This article expresses the author's views and does not necessarily reflect those of the Commission, the Commissioners, or other members of the staff.

I. Introduction

The 6% average annual outperformance of stocks over short duration bonds has attracted ample attention in the economics and finance literature. Because this outperformance was hard to explain using prevailing consumption-based asset pricing models, it is become known as the equity premium puzzle (Mehra and Prescott (1985) and Hansen and Singleton (1982)). The literature has proposed many different potential solutions since the introduction of the puzzle.[1] Theoretical explanations include richer consumption growth dynamics, sophisticated utility and belief specifications of the representative investor, and taxes. Another strand of the literature has questioned the puzzle's economic and statistical robustness.[2] This paper contributes to this latter strand and builds on the striking finding in Goetzmann and Jorion (1999) that U.S. equity performance appears to be an exception rather than the norm. These authors study a cross-section of countries and find that the longer a market survives, the higher its average realized rate of return, raising the possibility that survivorship bias is a contributor to the measured performance of markets still in existence. In this case, looking only at the best-performing country yields a biased estimate of the ex-ante expected return on equity.[3] In a similar spirit, Brown et al. (1995) model survival as the price level exceeding an absorbing lower bound. They show that simply conditioning on the price level never reaching the bound can address the equity premium puzzle documented in the U.S. equity market.

In this paper, we construct a comprehensive database with total return indices from a cross-section of 55 countries and revisit the question of whether survivorship bias plays an important role in the historical equity premium. While we show that equity has outperformed short-duration fixed income claims across countries in our sample, we also find that the

[1] See for example Campbell and Cochrane (1999), Bansal and Yaron (2004), Wachter (2013) and the references therein.

[2] See for example Cochrane (1997).

[3] Goetzmann and Jorion (1999) draw their conclusion by investigating the capital appreciation indices of 39 countries, but do not include dividend payouts. The latter constitutes an important fraction of realized returns, particularly in the early part of our sample.

U.S. appears lucky compared to most other countries. For example, the U.S. was never heavily implicated in either world war. The few crises that did take place, including the Cuban Missile Crisis, resolved themselves peacefully without straining the economy. As such, from the perspective of an early investor in the U.S. stock market it is impossible to predict with certainty the rise and ensuing dominance of the U.S. equity market. In fact, Argentina was among the world's ten richest countries early in our sample, with growth exceeding those of Canada and Australia. Unfortunately, its emergence was short-lived, and was soon followed by episodes of decline and stagnation. It is now well-known as a re-emerging market. With the benefit of hindsight, we know that it would have been unwise to invest in the Argentinean stock market at the turn of the century. Although investors may have expected a high return, they ended up earning a low realized return ex-post, despite high levels of volatility.[4] Overall, it motivates us to model crash risk via learning. By jointly studying a large cross-section of markets, we are able to trace the formation of subjective crash beliefs across time. By comparing the ex-post realized rates of return on each market against the ex-ante expected rates of return, we get a direct measure of the extent of survivorship bias.

To quantify the effect of survivorship bias on realized returns, we model an investor who learns about the frequency of stock market crashes from the cross-section of countries. Our definition of a market crash follows a simple threshold rule: an annual stock return below a certain cutoff level. While we choose to model learning about the frequency rather than the magnitude of disasters, our approach is easily adaptable for the severity of market crashes. In particular, while our analysis focuses on a threshold of -30%, we find similar results when using thresholds ranging between -20% and -35%.[5] It is also possible to estimate the subjective crash belief for a continuum of cutoffs and obtain a subjective return distribution

[4]When comparing returns across markets, it matters whether one is using geometric returns (GR) or arithmetic returns (AR). By Jensen's inequality, GR $\leq$ AR. If one assumes log-normality, AR = GR + $1/2\sigma^2$. Figure 1 plots the AR and GR against volatility in each market. Return and volatility are positively correlated for AR while negatively correlated for GR. Goetzmann and Jorion (1999) base their analysis primarily on geometric returns, although they note the same result holds for arithmetic returns.

[5]See Appendix B.

in terms of a kernel density function.[6] However, for ease of economic interpretation, we focus on one feature of the return data arguably of most interest to investors: downside crash risk. We improve upon the simple learning rule used by a naive investor, who only uses time series information. Instead, we impose a hierarchical structure on the problem by assuming that the crash risk of each country is drawn from the same distribution. The investor, in turn, has a hyperprior over the parameters governing the global crash risk distribution. The model allows the investor to learn about crashes in any particular country using the full cross-section of crashes across countries. This is achieved by periodically updating the parameters governing the global crash risk distribution. The strength of the cross-learning effect depends on the effective sample size of each country, as well as the uniqueness of the path of a country's return realizations, i.e., the extent to which it is an outlier. For a country like the US with a long and continuous price history, the extent of correction for survivorship bias is attenuated compared to a country like Estonia, an emerging equity market with only 23 years of history. Intuitively, it is not hard to see that for an infinitely long sample, we would optimally put all weight on country-specific observations when determining the frequency of rare events.

We estimate our model by applying a state-of-the-art Markov Chain Monte Carlo (MCMC) algorithm – Hamiltonian Monte Carlo (HMC) – to obtain the posterior crash belief of the investor. By replicating the information set of an investor at each point in time, we are able to track the subjective crash belief of an investor over time. Any Bayesian analysis necessarily requires the specification of a prior distribution at the start of the sample. We consider three such priors: an uninformative, a semi-informative and an informative prior and find that our results are robust to this choice. That is, as the sample progresses, the subjective crash belief for the U.S. shows a persistent and widening divergence from the implied global mean crash risk, particularly during the latter half of the 20th century. When evaluated within an equilibrium setting, the downward trend in the investor's subjective crash belief

[6] An alternative non-parametric approach is to model returns using the Dirichlet process. However, the resulting kernel density function of returns is harder to interpret economically.

(i.e., disaster risk) results in a declining equity premium and a rise in the valuation ratio of the stock market. Moreover, as investors learn about the data generating process, a lower probability of future crash risk results in positive return realizations, which is reflected in the 6% measured average excess return. We label the resulting wedge between the historical and expected risk premium as the *learning* wedge. In addition, with a well-defined posterior belief about disasters, we can evaluate the perception of *luck* in the U.S. equity market experience. We show that *luck* and *learning* jointly explain 2% out of the 6% historical equity premium, making the equity premium less of a puzzle, as pointed out by Avdis and Wachter (2017).

The paper draws inspiration from two strands of the literature on survivorship bias and crash risk. The heuristic known as survivorship bias refers to the tendency for people to focus on successful instances to draw inference about the underlying data generating process. Performance data of poorly-performing assets, with higher attrition rates, is often not readily available. Focusing attention on the surviving sample may lead researchers to overestimate performance. To avoid such bias in evaluating mutual fund performance, Jones and Shanken (2005) argue against the assumption of prior independence on the alphas realized by mutual funds. Instead, they discipline the inference about alphas by learning across funds.[7] In this paper, we apply Jones and Shanken (2005)'s approach to disciplining the variation of crash risk across countries. The rare-occurring nature of market crashes justifies the need for learning from the cross-section to discipline the estimate of country-specific crash risk. More importantly, it reveals to what extent the amount of luck in the U.S. experience can be attenuated by learning across countries.

Our paper also explores the idea that there are large negative potential shocks that did not realize in the U.S. sample, while investors are nonetheless compensated for the possibility of such events. To get a sense of the role played by such black-swan events, another strand of literature in finance focuses on the implied crash risk from other asset classes, in particular

[7] See also Stambaugh (2011).

out-of-the-money put options. The over-pricing of the options reveals that the underlying economy is far from Gaussian and highly negatively skewed. Nevertheless, it is not possible to back out the physical process from the risk neutral distribution implied from options without assumptions on risk preferences. Furthermore, the extent to which the equity premium puzzle can be attributed to rare events is also subject to debate (Backus et al. (2011), Welch (2016)). More recently, Dew-Becker et al. (2021) highlight the importance of jumps rather than forward-looking uncertainty for investors' marginal utility. In this regard, our paper provides a well-disciplined benchmark of physical crash belief formation, which allows us to learn from crashes taken place in other countries when inferring crash beliefs in the US.

Lastly, the paper contributes to a large literature of learning in financial markets. Baley and Veldkamp (2021) and Pastor and Veronesi (2009) provide a comprehensive review of the literature on learning in macro-finance. Kozlowski et al. (2020) model learning about downside risk in a non-parametric manner and show that the occurrence of a rare but salient event such as the great recession contributes to a persistent change in beliefs and the heightened awareness about tail risk helps explain the secular decline in interest rates. Nonetheless, their result relies on the agent not considering the possibility of a rare disaster before such event occurring. This is achieved via non-parametric learning, where there is an expansion in the sample space after realizing such event can in fact happen.[8] By contrast, our paper allows investors to entertain the possibility of such an event even without seeing any of such events occurring in-sample, and is therefore more conservative. In terms of asset pricing, Collin-Dufresne et al. (2016) model learning in an endowment economy and show that learning can generate endogenous risk premia and help resolve asset pricing puzzles. Our paper instead focuses on the contrast between ex-ante expected and ex-post realized

[8] In Bayesian learning, an interesting phenomenon is that there is a discontinuity between zero and any strictly positive number. For learning in a finite-dimensional setting, the posterior is absolutely continuous with respect to the prior. That is, the posterior distribution is defined over the same sample space as the prior. The posterior probability of an event can be non-zero only if the prior probability of the same event is non-zero. Otherwise, it is impossible to learn the possibility of such event in a Bayesian setting. However, the property can be violated in an infinite-dimensional setting, which is what non-parametric learning brings about.

rates of return and provides a normative benchmark to addressing survivorship bias in the U.S. equity market.

On the methodological side, the paper combines Bayesian filtering techniques with a hierarchical Bayesian model to infer latent states while teasing out country-level heterogeneity. In extensions to the model presented in the Appendix to this paper, we bring together the literature on regime switching models from Hamilton (1994), the change-point models introduced by Chib (1998) and later applied to infer breaks in the equity premium by Pástor and Stambaugh (2001), and the statistical literature on hierarchical Bayesian models advocated by Gelman et al. (2013). Finally, we estimate the model with advanced Monte Carlo sampling algorithms developed in Hoffman and Gelman (2011) and Neal (2011).

The paper proceeds as follows. In Section II, we first introduce the general setup of the model. We then discuss the choice of prior and derive the posteriors for different sets of parameters of interest. Section III discusses the data and methodology. In Section IV we present and discuss the results. Next, we explore the asset pricing implications of the secular decline in crash risk in the past century in Section V.

II. Model

In this section we introduce a parsimonious model in which investors learn about a latent but constant crash risk for each country, through the data observed in the cross-section of countries. We discuss extensions of the model to account for correlated market crashes in Appendix A, and show that the main conclusions from the paper are robust to this extension.

A. Correcting for Survivorship Bias

Prior. Consider an investor who has a common prior over the probability p_i of a stock market crash in country i, $i = 1, \dots, n$, given by:

$$p_i|\alpha, \beta \sim \text{Beta}(\alpha, \beta) \quad \forall i.$$

In our analysis, α and β are themselves uncertain. We will discuss the hyperprior on α and β in the next section, and for now simply refer to it as $f(\alpha, \beta)$.

Likelihood. Each period, the investor observes crash realizations in the cross-section of n countries. Crash realizations $D_{i,t}$ are independently and identically distributed across time and countries, conditioning on the country-specific crash probability (or crash risk) p_i. From the investor's point of view (which does not include knowledge of p_i) they will of course be correlated. The crash indicator $D_{i,t}$ takes a value of 1 if a crash occurs and 0 otherwise:

$$D_{i,t}|p_i \overset{i.i.d.}{\sim} \text{Bernoulli}(p_i).$$

The assumption that crashes are i.i.d. conditional on the parameters is made for convenience. Further enriching the model to enable it to identify a correlation structure would complicate the model and would not have first-order effect on inference regarding the unconditional crash probability for each country.[9] We nonetheless explore extensions to the model to account for correlated crash risk in Appendix A.

Let $Y_{i,\tau}$ be a vector of crash indicators of country i up to time τ:

$$Y_{i,\tau} = [D_{i,1}, D_{i,2}, \dots, D_{i,\tau}].$$

Posterior. The joint posterior distribution of the parameters of interest then follows from

[9] One possible second-order effect could come from the degree of precision in the estimation of α and β. To the extent that crashes are correlated, estimates of these parameters could be less precise given that there are fewer independent observations.

Bayes' rule:

$$
\begin{aligned}
f(\{p_i\}_{i=1}^n, \alpha, \beta | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) &\propto f(\{D_{j,t}\}_{j=1,t=1}^{n,\tau} | \{p_i\}_{i=1}^n, \alpha, \beta) f(\{p_i\}_{i=1}^n | \alpha, \beta) f(\alpha, \beta) \\
&= \prod_{i=1}^n p_i^{Y_{i,\tau}\mathbf{1}^\top} (1-p_i)^{\tau - Y_{i,\tau}\mathbf{1}^\top} \prod_{i=1}^n p_i^{\alpha-1}(1-p_i)^{\beta-1} f(\alpha,\beta) \\
&= \prod_{i=1}^n p_i^{Y_{i,\tau}\mathbf{1}^\top + \alpha - 1} (1-p_i)^{\tau - Y_{i,\tau}\mathbf{1}^\top + \beta - 1} f(\alpha, \beta)
\end{aligned} \tag{1}
$$

Conditional upon knowing the parameters that govern the global crash risk distribution (α and β), the investor infers the crash risk in each country i only from country-specific information, that is, $f(p_i|\alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = f(p_i|\alpha, \beta, Y_{i,\tau})$. It follows from (1) that

$$
p_i | \alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau} \sim \text{Beta}(Y_{i,\tau}\mathbf{1}^\top + \alpha, \tau - Y_{i,\tau}\mathbf{1}^\top + \beta) \quad \forall i \tag{2}
$$

The conditional posterior (2) has a convenient interpretation. At each point in time τ, the investor combines the count of crashes in the country of interest (country i) with a prior pseudo-sample consisting of α crashes out of a total of $\alpha + \beta$ observations. At each point in time, the full data informs this prior pseudo-sample.

While (2) conditions on specific values of α and β, these themselves update dynamically over time. The laws of conditional probability imply

$$
f(\alpha, \beta | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = \frac{f(\{p_i\}_{i=1}^n, \alpha, \beta | \{D_{j,t}\}_{j=1,t=1}^{n,\tau})}{f(\{p_i\}_{i=1}^n | \alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau})}. \tag{3}
$$

To arrive at the posterior distribution for the crash probability given the data, the investor needs to integrate out the uncertainty over the global crash risk parameters:

$$
f(p_i | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = \int_{\alpha,\beta} f(p_i | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \alpha, \beta) f(\alpha, \beta | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) d(\alpha, \beta) \quad \forall i. \tag{4}
$$

Equation 3 and 4 highlight how the model is able to achieve cross-learning. The model uses all the information from the cross-section up till time τ to derive the joint posterior of

α and β, as shown in Equation 3. At any point in time, the investor forms a posterior for the global mean crash risk (equivalently, crash probability of an average country), $\alpha/(\alpha+\beta)$. This posterior can in turn be viewed as a prior on the crash risk for any specific country, as we show below. The effective sample size, $\alpha+\beta$, controls the spread of the Beta distribution, with the dispersion in global crash risk decreasing in $\alpha+\beta$.

The global crash risk parameters (α and β) further feed into the estimation of the country-specific crash risk in Equation 4. After we integrate out uncertainty with respect to these global crash risk parameters, we arrive at the marginal probability distribution for country-specific crashes. The first component of the integrand is the conditional posterior of the crash probability for country i shown in Equation 2, which peaks at$(Y_{i,\tau}\mathbf{1}^\top+\alpha)/(\alpha+\beta+\tau)$. This is the posterior mean one would arrive at by Bayesian updating according to the common prior and the observation that $Y_{i,\tau}\mathbf{1}^\top$ crashes occurred in τ periods. Alternatively, one may write the posterior mean of country-specific crash risk p_i as a weighted average:

$$\mathbb{E}(p_i|\alpha,\beta,\{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = \frac{Y_{i,\tau}\mathbf{1}^\top}{\tau}\frac{\tau}{\tau+\alpha+\beta}+\frac{\alpha}{\alpha+\beta}(1-\frac{\tau}{\tau+\alpha+\beta}) \tag{5}$$

The subjective crash risk p_i increases in the number of crashes in country i observed up till time τ, $Y_{i,\tau}\mathbf{1}^\top$. It also increases in the global mean crash risk, $\alpha/(\alpha+\beta)$. More interestingly, as the global crash risk distribution becomes more concentrated, i.e. when $\alpha+\beta$ increases, the country-specific crash risk p_i is shrunk towards the global mean crash risk, $\alpha/(\alpha+\beta)$:

$$\frac{\partial\,\mathbb{E}[p_i|\alpha,\beta,\{D_{j,t}\}_{j=1,t=1}^{n,\tau}]}{\partial(\alpha+\beta)} = \frac{-Y_{i,\tau}\mathbf{1}^\top+\frac{\alpha}{\alpha+\beta}\tau}{(\alpha+\beta)^2(1+\frac{\tau}{\alpha+\beta})^2}\begin{cases}\geq 0 & \text{if } \frac{Y_{i,\tau}\mathbf{1}^\top}{\tau}\leq\frac{\alpha}{\alpha+\beta}\\ <0 & \text{if } \frac{Y_{i,\tau}\mathbf{1}^\top}{\tau}>\frac{\alpha}{\alpha+\beta}.\end{cases} \tag{6}$$

Thus, when the global crash risk is estimated precisely, posterior means on riskier countries are scaled down while posterior means on less risky countries are scaled up. The hierarchical Bayesian model thus achieves a shrinkage effect: the use of all the data imply that country-specific crash risk is shrunk towards the common mean.

The strength of the shrinkage effect depends on the effective sample size of country i. Suppose that countries differ in sample length τ_i, one may write the conditional posterior mean crash risk in the following alternative form and look at the asymptotic behavior.

$$\lim_{\tau_i \to \infty} \mathbb{E}(p_i|\alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = \lim_{\tau_i \to \infty} \frac{\frac{Y_{i,\tau_i}\mathbf{1}^\top}{\tau_i} + \frac{\alpha}{\tau_i}}{1 + \frac{\alpha+\beta}{\tau_i}} = \lim_{\tau_i \to \infty} \frac{Y_{i,\tau_i}\mathbf{1}^\top}{\tau_i}$$

As the sample size τ_i increases towards infinity, the conditional posterior mean crash risk converges to the true crash risk in the country by the law of large numbers. It implies that the hyperprior has a diminished effect on the estimated crash risk in country i if the market has a long history. In other words, the extent of the shrinkage effect, as a means to deal with a small sample inference problem, is therefore data-driven.

Finally, the marginal posterior of the country-specific crash risk p_i integrates out uncertainty with respect to the global crash risk distribution, as in Equation 4. The resulting mean country-specific crash risk becomes a weighted average of the conditional means if global crash risk is observed. That is,

$$\mathbb{E}[p_i|\{D_{j,t}\}_{j=1,t=1}^{n,\tau}] = \mathbb{E}[\mathbb{E}(p_i|\alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau})|\{D_{j,t}\}_{j=1,t=1}^{n,\tau}].$$

We conclude this section by emphasizing that the simple hierarchical structure of the model imposes innocuous assumptions on the investor's subjective belief while allowing her to aggregate information from the cross-section of countries to form beliefs about the crash risk in the US.

In the parsimonious model presented so far, we assume conditional independence of market crashes across countries. In Appendix A, we enrich the model for correlated market crashes, and also discuss why modeling time-varying crash risk is not quite feasible given limited information. In the next section, we explain the our prior choices in further detail.

B. *Hyperpriors*

Recall that our common prior on the country-specific crash risk follows a Beta distribution: Beta(α, β). Consider the following the change of variables:

$$\phi = \frac{\alpha}{\alpha + \beta}, \quad \lambda = \alpha + \beta.$$

This change of variables facilitates interpretation of the parameters. That is, the mean of the Beta distribution ϕ is the global average crash risk. The effective sample size λ controls the dispersion of the distribution: the distribution of crash risk is more dispersed with a low λ. We choose the hyperprior on λ to have a relatively long tail, implicitly allowing the global crash risk distribution to be dispersed.

We consider a range of hyperpriors. To arrive at our uninformative hyperprior, we follow Gelman et al. (2013) in assuming that ϕ is uniformly distributed between 0 and 1. We model λ to have an effective sample size of at least one and follows a Pareto distribution. Overall, the prior on the two parameters can then be summarized as follows:

$$\phi \sim \text{Uniform}[0, 1],$$
$$\lambda \sim \text{Pareto}(1, 0.5).$$

Hence, the joint prior distribution for ϕ and λ is

$$f(\phi, \lambda) \propto \lambda^{-3/2} \mathbf{1}_{\phi \in [0,1]}.$$

The uninformative hyperprior on α and β is therefore

$$f(\alpha, \beta) = \begin{vmatrix} \frac{\partial \phi}{\partial \alpha} & \frac{\partial \phi}{\partial \beta} \\ \frac{\partial \lambda}{\partial \alpha} & \frac{\partial \lambda}{\partial \beta} \end{vmatrix} f(\phi, \lambda) = (\alpha + \beta)^{-5/2}.$$

While uninformative priors have theoretical appeal, they do imply assumptions that may seem extreme: in particular that the mean global crash risk is 50%. This is one reason to also consider two additional hyperprior specifications: a semi-informative and an informative prior. As our *semi-informative* hyperprior we consider the following distributions for ϕ and λ:

$$\phi \sim \text{Uniform}[0, 0.35], \quad \lambda \sim \text{Pareto}(1, 2.5).$$

In this case, the average crash risk is assumed to be uniformly distributed across the range of the in-sample frequency of market crashes across countries. That is, we use ex-post information to discipline the prior, which provides a moderate restraint on the values that the probability of market crashes can take on. We also change the shape parameter of the Pareto distribution for λ to 2.5, so that the distribution has a finite variance.

As our *informative* hyperprior, we consider:

$$\phi \sim \text{Beta}(2, 98), \quad \lambda \sim \text{Pareto}(1, 2.5).$$

In this case, the average crash risk has a mean of 2%. This is motivated by the fact that there is one market crash in the U.S. time series during the 50-year period from 1870 to 1920, using Robert Shiller's data. As was the case for the semi-informative hyperprior, we change the shape parameter of the Pareto distribution to 2.5, so that the distribution has a finite variance. In using priors that are informative and motivated by the data, but not equal to the eventual posterior, we follow Cogley and Sargent (2008) and Farmer et al. (2023).

The top row of Figure 2 compares the sampling distributions of the idiosyncratic crash risk for all three hyperpriors. For both the informative and the semi-informative cases, the hyperprior is heavily right-skewed and does not depict the bimodal pattern as in the uninformative hyperprior. The informative hyperprior has a global crash risk that centers around 2% with a cross-sectional variation of 0.8%. From the figure, most countries are assumed to have a crash risk lower than 2%, though some countries in the thin right tail have larger crash risk. Similarly, the semi-informative hyperprior has a mean global crash risk of 17% with a cross-sectional variation of 6%, with more countries having higher crash risk compared to the informative case. We present our findings for all three specifications and show that the results are robust to the choice of the hyperprior. This is because as long as the hyperprior is sufficiently diffuse, the data will drive most of the movement in the posterior belief.

III. Data and Methodology

A. Data

Data on annual stock returns come from Global Financial Data (GFD), which provides long time series of stock return indices for global equity markets. Our sample starts in 1920 and ends in March 2020, spanning a century. The choice of starting time is influenced by data availability.[10]

Our sample covers a total of 55 countries. This includes all the MSCI developed markets, almost all the emerging markets, and several frontier markets. The main series we use are the total return indices, which represent the cumulative buy-and-hold return from investing in the index, with reinvestment of dividends.[11] Since such series are unavailable for most countries in our sample, we construct total return indices ourselves using index price levels combined with information on monthly reinvested dividends. All returns are in real dollar terms, in order to emulate the rate of return a U.S. investor would earn by investing in global equity markets. The tickers used to retrieve data from GFD are in Table III, and the summary statistics are provided in Table I.

[10]Some suggest that the 1920s marked a change in how Americans thought about the stock market, perhaps influenced by 20% inflation during World War I. Realizing that the U.S. dollar no longer retains its value in bank deposits, investors showed growing interests in equity investments. See, e.g., `https://www.forbes.com/sites/briandomitrovic/2020/01/09/why-did-people-buy-stocks-in-the-1920s/?sh=1d8b4cf03390`.

[11]See Figure 3 for the cumulative buy-and-hold returns across countries in our sample.

Market segmentation is substantial for most of the 20th century, which implies that a U.S. investor cannot freely invest in foreign equity markets. The literature in international finance disagrees on the appropriate numeraire in making cross-country comparisons (Dimson et al. (2002)). The real exchange rate is shown to be stable in the long run but not the short run. Despite the failure of absolute purchasing power parity, the notion of relative purchasing power parity, which states that countries with high inflation should see their currency depreciating, is accepted more broadly. Since we focus on the belief formation of a representative U.S. investor, rather than the actual implementation of such trading strategies, we believe the usage of a common currency denomination is appropriate for our purposes. More importantly, we can derive an estimate of the equity premium in each country by comparing its average real dollar return against the U.S. treasury rate.

At the start of our sample, not every country has an active stock exchange. We allow markets to enter through time, implying that we have an unbalanced panel with an expanding cross-section of countries. We aim to be as extensive as possible in our data collection, but we acknowledge that there are a few existing markets for which we do not have data. Examples are Argentina prior to 1946 and Brazil prior to 1953. In such cases, data are more likely to be missing in bad times than good times (and for badly-functioning as opposed to well-functioning markets) – which is to say that even our sample is not completely free of survivor bias. Hence, the missing data contain potentially a higher crash frequency than what we observe in-sample. Since the U.S. crash risk is our primary focus, and there is little doubt about the reliability of the U.S. data, the selection in other markets can only bias downward our survivorship-corrected U.S. crash risk estimate. Table II outlines the coverage of our data, and the frequency of market crashes in each country.

Moreover, countries exit and re-enter the sample throughout time. In international equity markets, unlike the US, not every country has a continuous price history. In fact, almost half of the markets in our sample experience some forms of hiatus in their trading (see Table II). Such breaks can either be temporary or may stretch for decades. For instance, most

European countries have their stock exchanges shut down during World War II, usually for a couple of months. The Spanish Civil War prevents Spain from being in the sample for four years. As for most countries that have/had Communist regimes, investors lost the entirety of their investments when nationalizations took place. In all of these cases, the returns near the break points are of particular interest to us, as they represent potentially disastrous returns to investors.

It is important to highlight the difference in our treatment of the data from the method taken by Goetzmann and Jorion (1999). These authors also distinguish between two types of disruptions in the equity market: temporary and permanent closures. Returns are interpolated in periods where there is a temporary closure of the stock exchange. Since our observation is at an annual frequency, temporary closures lasting less than a year are of little relevance to us. In contrast to our approach, they assume a drop of -75% in the case of a permanent closure, which took place in the aforementioned countries that experienced communist regimes. Bialkowski and Ronn (2016) argue that the residual value of equity following major socio-economic disruptions are overstated in Goetzmann and Jorion (1999). Instead, they argue that severe restrictions in civil and property rights, resulting in the inability to consume freely out of one's wealth, is equivalent to a collapse in the capital market. We follow their approach[12] in assuming a -100% return for stock markets in years when such a restriction is imposed, and re-include these markets in our sample once basic rights in their respective countries are reinstalled. Such adjustments are made for countries heavily implicated in World War II, for example Japan, France, and Germany. We believe the approach is reasonable because the capital and price controls in war time render the return series unrepresentative of the actual return earned by investors.

Given our goal of studying crash risk, we identify a market crash as an annual return below -30%.[13] It is the frequency of such an event that the investor in our model is learning about and correcting survivorship bias for. We choose not to include learning about the

[12]See Table 7 in Bialkowski and Ronn (2016).

[13]We show robustness of our results for cutoff -20% and -35% in Appendix B.

magnitude of the crash in our model, because simply raising the bar of classifying a market crash can inform the investor of the incremental subjective likelihood she assigns to a slightly more severe market crash. Presumably, one may push the model to the limit and estimate the posterior belief for any small interval of return. This would essentially yield a semi-parametric estimate of the return expectation by aggregating information from the cross-section.

B. Method

We construct the posterior using Hamiltonian Monte Carlo (Neal, 2011). To implement the sampler, we start four Markov chains at randomly selected starting values and draw a sample of length 10,000 from each chain[14], where the first 2,000 draws serve as burn-ins.We then pool the draws across chains for the posterior distribution. The exercise is repeated with an expanding window through time, representing the accumulation of information the investor uses to update her crash belief. We apply the standard tests diagnostics to compare within and between chain variation to ensure the convergence of the algorithm.[15] to ensure the convergence of the algorithm.

IV. Model Results

A. Uninformative Prior

Cross-Sectional Comparison. In Figure 4, we plot the subjective crash risk estimate (red dots) against the in-sample frequency of market crashes (gray bars) for the uninformative prior. We can immediately see the shrinkage effect induced by the hierarchical model. For the United States, the shrinkage effect is moderate: it increases the crash risk from 5% (in-sample) to 6.54% (posterior mean). That is, the investor would correct the survivorship

[14] A sample size of 10,000 might seem small in standard MCMC exercises. However, it is more than sufficient for HMC (Hoffman and Gelman (2011), Betancourt et al. (2017)). In fact, it is due to the ability of the HMC to move around quickly with its gradient-based algorithm that we opt for the approach.

[15] Such tests include trace plots, checking the effective sample size, and applying the Gelman and Rubin (1992) diagnostics to compare within and between chain variation.

bias in terms of crash risk in the U.S. sample by 30.8%. The strongest shrinkage effect is that for Turkey: the posterior mean crash risk is shrunk by almost half. Meanwhile, two countries stand out in having no market crashes worse than -30% in history: Switzerland and Morocco. Nevertheless, the posterior mean crash risk differs between these two countries. For Switzerland, which we have 100 years of data for, the shrinkage toward the mean effect is weaker than that for Morocco, an emerging stock market. Despite the fact that both countries are significant outliers in having safe returns overall, the model borrows more information from the cross-section to discipline the perceived crash risk for Morocco because of a lack of sufficient observations. As a result, Morocco is deemed to be a riskier country than Switzerland, and its good performance is more likely to be attributed to luck.

Another notable pattern from the figure is that the subjective crash risk is not monotonic in the historical frequency of market crashes. This is due to the updating scheme of the investor. As discussed before, the cross-learning effect is stronger for markets with smaller effective sample sizes and hence shorter histories. The investor rationally compensates for the lack of information by "borrowing" more information from the cross-section. The periods in which market crashes take place also matter for the subjective crash risk. This is due to the fact that the global crash risk estimate is inferred from the country-specific crash risk estimates. Suppose that in an especially risky period, markets crashed in many countries. The first order effect is that there is an increase in the country-specific crash risk estimate. In addition, the mode of the global crash risk distribution is likely to shift to the right. This further increases the subjective crash risk in all markets, even those without a crash, by Equation 5. What's more, suppose that the global crash risk is estimated more precisely as a result ($\alpha + \beta$ increases), then the shrinkage towards the global mean is stronger by Equation 6.

What is the average market in the sample? Figure 4 reveals that it is Czech Republic, of which the subjective crash risk and the in-sample frequency are both around 9.1%. For the average market, there is no shrinkage, while crash risk in other markets are all shrunk towards this number. However, we caution that the identity of this average market is time-varying. Although Hong Kong is the average market in terms of crash risk in 2020, this is not the case in the 1950s, since its stock market did not exist back then.

U.S. v.s. An Average Country. Figure 5 plots the subjective crash risk for the U.S. and the average country (global mean crash risk) from 1921 to 2020. The upper panel plots the posterior means while the lower panel contains the 95% credible set as well. The credible set is wide and asymmetric. While the credible set starts out as wide and highly right-skewed, it quickly shrinks and becomes more symmetric, though an aspect of right-skewness remains. Note that the credible set for the global crash risk parameters shrinks most dramatically, such that the U.S. mean actually lies outside of this set toward the end of the sample.

In particular, we observe that the global mean crash risk quickly stabilizes around 8.75%. The abundance of data pooled from the cross-section allows a more precise estimate of the global mean crash risk compared to the country-specific ones. The subjective crash risk in the U.S. initially departs from the global mean before spiking to a similar level in 1931, when the Great Depression hits. The 1937 crash boosts the subjective crash belief in the U.S. further above that of the global mean for a brief period of time. This is due to the fact that the 1937 crash is rather unique to the US. Afterwards, the subjective crash risk in the U.S. departs steadily away from that of the global mean, due to the absence of negative shocks in the U.S. sample. We observe that the spikes in subjective crash risk in 1974, 2008 and 2020 move in tandem with those in the global mean crash risk, as they are all systematic shocks. Combined with the fact that the posterior estimate's variance is shrinking, the adjustments in subjective crash belief are attenuated. As of March 2020, there is a 2.62% difference between the crash risk in the U.S. (6.54%) and the global average (9.16%).

B. Alternative Hyperprior Specifications

We reestimate the model with the two alternative hyperprior specifications introduced in Section II.B, and present the results in Figure 2. Overall, the inferred crash risk in all three specifications show similar comovement, though with different starting values. Notably, the figures highlight the role of the hyperprior in influencing the investors' belief at the beginning of the sample. However, the influence wears down overtime and the beliefs all converge to a similar value in 2020.

The investor with an informative prior stays optimistic till the Great Depression, when the inferred crash risk for both the global mean and the U.S. spikes to around 8% from near 2%. From the point onwards, the global mean crash risk continues its drift upwards, implying that the hyperprior is indeed too optimistic. The U.S. crash risk continues its drift downwards, similar to the results of the uninformative prior. Overall though, the belief paths lie slightly below those for the semi-informative and uninformative case, reflecting the effect of the informative hyperprior.

For the semi-informative case, the belief evolution for the investor is very similar to that of an agent with an uninformative prior. The hyperprior only affects the inferred crash risk in the beginning of the sample. From 1930 onwards, the belief evolutions are essentially identical to those in our uninformative prior specification.

Overall, we find that the inferred crash risk across the three specifications are very similar to one another. The prior is effective in the early periods of the sample, but quickly wears off. To avoid volatile belief movements in the beginning of the sample that is largely due to the diffuse hyperprior, we use the first 10 years of the sample to initiate our prior and anchor investor's belief. We zoom in on the slow downward drifts in U.S. crash risk from 1931 to 2020 to investigate the main asset pricing implications.

V. Asset Pricing Implications

In order to assess the asset pricing implications for the secular decline in crash risk, we first map the partial equilibrium concept of crash risk to its general equilibrium analog of a consumption disaster. In the disaster risk literature, pioneered by Rietz (1988) and Barro

(2006), and later extended in Wachter (2013) and Gabaix (2012), the equity premium puzzle can be resolved by the presence of a low-probability disastrous outcome in consumption growth. Investors demand a high risk premium to compensate for the possibility that equity pays off poorly in states where marginal utility is high.

To draw a parallel between crash risk and disaster risk, we make use of the finding in Nakamura et al. (2013) that the stock market crashes at the onset of disasters because the agent rationally takes into account the duration of the disaster and the subsequent recovery. In order to extend the beliefs about crash risk to the beliefs about consumption disasters, we assume they follow the same trend.[16]

Due to the rare occurrence of consumption disasters in-sample, we depart from the rational expectation equilibrium in assuming that the agent does not know the true extent and evolution of disaster risk. Instead, the agent prices assets with the subjective expectation about disaster risk, which maps to crash risk.

As is standard in the disaster risk literature, consumption follows a random walk with drift and jump. Log consumption growth has mean μ and standard deviation σ in periods without disasters.

$$\Delta c_{t+1} = \mu + \sigma \epsilon_{t+1} + v_{t+1}$$

Our setup departs from the standard one in that the disaster risk p is no longer a known constant. Instead Bayesian learning produces a series of probability density functions that reflect the investor's uncertainty toward the disaster risk parameter p. To predict the possibility of a disaster next period, the investor integrates out uncertainty with respect to p. The posterior mean p_t serves as a sufficient statistic for the disaster risk next period.[17]

[16] An alternative explanation for the same trend assumption is that market crashes serve as signals to consumption disasters, as they contain forward-looking expectations about macro fundamentals. With Bayesian updating, the subjective disaster risk will also trend downwards.

[17] Note that the posterior mean is a sufficient statistic if the agent has a Beta prior and observes crash realizations that follow a Bernoulli distribution. The predictive density $p(v_{t+1} = 1) = \int p(v_{t+1} = 1|p) f(p|I_t) dp \propto \int p \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} dp = \frac{\alpha}{\alpha+\beta}$. This is not exactly true in our setup, because we allow the investor

Consumption then drops by b upon impact.

$$v_{t+1} = \begin{cases} \log(1-b) \quad \text{w.p. } p_t \\ \\ 0 \quad \text{w.p. } 1-p_t \end{cases}$$

The representative consumer has an Epstein-Zin type utility function, where γ is the degree of relative risk aversion, ψ is the willingness to substitute across time, and δ is the rate of time preference.

$$V_t = \left[(1-\delta)C_t^{1-\frac{1}{\psi}} + \delta\,\mathbb{E}(V_{t+1}^{1-\gamma})^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}$$

When $\gamma = 1/\psi$, the utility function boils down to the standard CRRA form. We start with the CRRA case to discuss the asset pricing results, and illustrate why we need Epstein-Zin for some of the asset pricing results to hold.

A. *Declining Equity Premium*

First, the secular decline in disaster risk implies a declining equity premium in the past century. A learning agent demands a lower risk premium to invest in the U.S. equity market as she increasingly believes that it is a safe place to invest in.

With CRRA utility, the equity premium on the consumption claim increases in the subjective disaster risk p_t for a risk averse agent.

$$\log \frac{E_t(1+R_{t+1})}{1+R_f} = \gamma\sigma^2 + b[(1-b)^{-\gamma} - 1]p_t \tag{7}$$

The decline in the equity premium has been widely noted in the literature, e.g. Lettau et al. (2008). This decline contributes to the upward trend in the market valuation ratio in the past century, although the change in dividend payout policy and the secular decline

a common prior over global crash risk to correct for survivorship. However, with enough data and to a first-order approximation, the Beta-Bernoulli conjugacy result implies that the posterior mean crash risk is the subjective probability that a market would crash in the future.

in interest rates may also have played a role.[18] By contrast, the stability of global crash risk implies that there is a growing divergence between the risk compensation in the global equity market compared to that of the U.S. market.

B. Decomposition of the Equity Premium

With imperfect information, the agent who learns about disaster risk would experience a series of permanent shocks to her beliefs.[19] In turn, the belief updating would result in a sequence of discount rate and cash flow shocks that manifest themselves in the historical returns. We can therefore quantify the upward bias in the U.S. historical realized equity premium that is due to learning-induced positive return news and the negative potential shocks that did not materialize in the sample studied. Updating p_t leads to a change in the valuation ratio, In order to focus on the econometric measurement of luck and learning, this exercise assumes the agent does not anticipate future adjustments in the parameters when in pricing assets.[20]

Suppose that the agent treats the current posterior mean as the true constant parameter and ignores any dynamics in the endogenous state transition that arise from learning, the valuation ratio is now time varying only from updating. That is, the price-consumption ratio experiences a permanent change each period because the agent arrives at a new estimate for the disaster risk p_t. However conditioning on the current information set, the investor believes the price-consumption ratio will be constant going forward.

The price-consumption ratio, $Z_t^C = P_t^C / C_t$, in the economy depends on p_t. Let K_t be a

[18]See van Binsbergen (2020) for an analysis of stock market performance relative to duration-matched fixed income instruments.

[19]The posterior means form a martingale, which follows form the law of iterated expectation: $E[p|I_t] = E[E(p|I_{t+1})|I_t]$

[20] This is otherwise known as anticipated utility, and was introduced by Kreps (1997) and used by Cogley and Sargent (2008). (Weitzman, 2007) incorporates the entire posterior distribution in a simpler learning problem, but, as in the present case, does not take future learning into account. Wachter and Zhu (2023) consider learning about a latent state, but assume other parameters are known.

monotonic transformation of the price-consumption ratio Z_t^C, we have that

$$K_t \equiv \frac{Z_t^C}{1+Z_t^C} = \beta e^{(1-\gamma)\mu+\frac{1}{2}(1-\gamma)^2\sigma^2}[(1-b)^{1-\gamma}p_t + 1 - p_t].$$

The expected return satisfies

$$\mathbb{E}_t[R_{t+1}] = \mathbb{E}_t\left[\frac{P_{t+1}^C + C_{t+1}}{P_t^C}\right] = \mathbb{E}_t\left[\frac{Z_t^C+1}{Z_t^C}\frac{C_{t+1}}{C_t}\right].$$

Note that the second equality makes use of the assumption that the agent treats the current posterior mean disaster risk as the true disaster risk going forward. As a result, the forward-looking valuation ratio is constant based on time t information. By contrast, the ex-post realized return is

$$R_{t+1} = \frac{P_{t+1}^C + C_{t+1}}{P_t^C} = \frac{Z_{t+1}^C+1}{Z_t^C}\frac{C_{t+1}}{C_t}.$$

The realized return reflects changes in the valuation ratio that are not taken into account from an ex-ante perspective. The difference between the expected and realized returns arises from the change in the valuation ratio as well as cash flow surprises. For a sample in which disasters did not materialize, such as the experience in the US, the average cash flow realization, $\frac{C_{t+1}}{C_t}$, is approximately $e^{\mu+\frac{1}{2}\sigma^2}$, higher than the expected growth in cash flows, $e^{\mu+\frac{1}{2}\sigma^2}[(1-b)p+1-p]$. In addition, if the valuation ratio rises following a decline in disaster risk, the realized return would be higher than the expected return. Note, however, that in order for the valuation ratio to go up when disaster risk declines (the economically intuitive direction), we need that $\gamma < 1$. Otherwise, the rise in the risk free rate will more than offset the decline in the risk premium, leading to a decline in the valuation ratio.[21] If the agent is endowed with a different willingness to smooth across time and across states, the rise in the risk-free rate following a decline in disaster risk can be attenuated and the return surprises

[21] This is a problem widely noted in the disaster risk literature since Barro (2006).

would be uniformly positive.

To illustrate this point, we price assets using a model featuring Epstein-Zin preferences. To capture the quantitative magnitude of the upward bias in the equity premium, we model a dividend claim as a levered claim on consumption, i.e. $\Delta d = \lambda \Delta c$.

Let $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$. The pricing kernel in the economy is affine in log consumption growth Δc_{t+1} and the return on the market portfolio $r_{m,t+1}$.

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1) r_{m,t+1}.$$

By the Euler equation, we have the pricing equation:

$$P_t^D = \mathbb{E}_t[exp(m_{t+1})(P_{t+1}^D + D_{t+1})].$$

The equilibrium valuation ratios, $Z_t^D = P_t^D / D_t$, are again functions of the current subjective disaster risk. Let K_t^D be a monotonic transformation of the price-dividend ratio:

$$K_t^D \equiv \frac{Z_t^D}{Z_t^D + 1} = \delta e^{(\lambda - \frac{1}{\psi})\mu + \frac{1}{2}[(1-\gamma)(\gamma - \frac{1}{\psi}) + (\lambda - \gamma)^2]\sigma^2} [(1-b)^{1-\gamma} p_t + 1 - p_t]^{\frac{\gamma - \frac{1}{\psi}}{1-\gamma}} [(1-b)^{\lambda - \gamma} p_t + 1 - p_t].$$

For the valuation ratio to be non-increasing in disaster risk, we need that $Z_{t+1}^D > Z_t^D$, or equivalently $K_{t+1}^D > K_t^D$, when $p_{t+1} < p_t$. That is,

$$\frac{dK_t^D}{dp_t} = \frac{d\big[[(1-b)^{1-\gamma} p_t + 1 - p_t]^{\frac{\gamma - \frac{1}{\psi}}{1-\gamma}} [(1-b)^{\lambda - \gamma} p_t + 1 - p_t]\big]}{dp_t} < 0.$$

A sufficient condition for this to hold is that $\lambda > \gamma \geq \frac{1}{\psi}$.[22] The leverage ratio on the dividend claim needs to be greater than the risk aversion and the agent weakly prefers early resolution of uncertainty, which is a standard calibration for the stock market.

Comparing the realized and the expected rates of return, we see that the difference mainly

[22]For a detailed discussion on the necessary and sufficient conditions, see Tsai and Wachter (2015).

arises from two channels: *luck* and *learning*.

$$R^D_{t+1} - \mathbb{E}_t[R^D_{t+1}] = \underbrace{\mathbb{E}_t\left[\frac{D_{t+1}}{D_t}\right]\left(\frac{Z^D_{t+1}+1}{Z^D_t} - \frac{Z^D_t+1}{Z^D_t}\right)}_{Learning} + \underbrace{\frac{Z^D_t+1}{Z^D_t}\left(\frac{D_{t+1}}{D_t} - \mathbb{E}_t\left[\frac{D_{t+1}}{D_t}\right]\right)}_{Luck}$$
$$+ \underbrace{\left(\frac{Z^D_{t+1}+1}{Z^D_t} - \frac{Z^D_t+1}{Z^D_t}\right)\left(\frac{D_{t+1}}{D_t} - \mathbb{E}_t\left[\frac{D_{t+1}}{D_t}\right]\right)}_{Residual} \tag{8}$$

Luck in the realized excess returns is approximately equal to the difference between the realized dividend growth and the expected dividend growth, $\frac{D_{t+1}}{D_t} - \mathbb{E}_t[\frac{D_{t+1}}{D_t}]$, as $\frac{Z^D_t+1}{Z^D_t}$ is close to 1. The term captures how much of the historical equity premium is biased upwards due to disasters that did not occur in-sample.

In addition, as the agent revises her perception of disaster risk downwards over the century, the ex-post price-dividend ratio similarly experiences revisions upwards. Although the agent believes ex ante that the price-dividend ratio will be constant going forward, the decline in disaster risk produces positive return news, again creating a wedge between the realized and expected rate of returns. We attribute the resulting upward bias in the historical equity premium to *learning* as the residual term is negligible in magnitude.

We calibrate the model using standard parameters from the disaster risk literature.[23] We assume that the magnitude of the decline in consumption equals that of the stock market crash, which is consistent with the full-information version of this model in which the distribution of consumption growth, given knowledge of the parameters, is i.i.d.[24] We feed in the posterior mean crash risk from the Bayesian estimation over time. We compute the expected stock returns for each of the 55 markets, with the probability of a disaster taking place at each point in time equal to the current posterior mean crash belief. We compare them against

[23] That is, $\gamma = 3$, $\beta = 0.98$, $\mu = 0.0252$, $\sigma = 0.02$, $\lambda = 2.6$, $\psi = 1$, and $b = 0.3$.

[24] Constantinides (2008) and Julliard and Ghosh (2012) note that declines of this magnitude are spread out over multiple years. Fixing preference parameters, a greater probability of a smaller disaster leads to a lower equity premium due to Jensen's inequality. However, if one accounts explicitly for multiperiod disasters, in a model in which agents have a preference for early resolution of uncertainty, the results are largely isomorphic to one in which the consumption decline occurs all at once (Tsai and Wachter, 2015). In such a model, the market crash occurs upon the onset of the multi-period disaster.

the model implied realized returns, with the realization of disasters corresponding to what happened in-sample. To ascertain that the model captures actual movements in stock prices, we compute, for each country, the correlation between the model implied realized returns and the realized stock returns in the data. Averaging across countries, such correlation is 50% over the 100-year period, suggesting that the parsimonious model correctly captures to some degree the actual movement in stock returns.

In Figure 6, we decompose the wedge between realized and expected rates of return[25] into the aforementioned *luck* and *learning* components. We compute luck and learning for data beginning in 1931, after the agent has seen ten years of data. For the United States, the realized rate of return exceeds the expected rate of return by 2%, which can be interpreted as the amount the historical equity premium has been overstated by. This 2% is about equally split between *luck* and *learning*. Interestingly, although our analysis indicates that the historical equity premia in Switzerland and Morocco have the same level of upward bias, Switzerland is shown to be less lucky while the majority of the sample bias is due to learning. This is because the investor in the model decides that Switzerland is a low-risk country early on. The decision attenuates the perception of luck and contributes to a steady downward revision in disaster risk. By contrast, the investor in the model ascribes most of Morocco's excess return to *luck* because it is deemed to be riskier than Switzerland. Crucially, the results do not depend on the choice of hyperprior (see Figure 2 and Table IV).

We can also use asset prices to evaluate whether our model is a plausible account of investor expectations. To do this, we can feed our posterior crash risk probability through the asset pricing model, together with observed crashes, and see if the resultant returns match that of the data. This is a difficult hurdle for our stylized model to meet in that it is only designed to match crashes and changes in long-run crash probabilities.

In Figure 7, we plot the 10-year rolling average of realized returns in the model against

[25]Here the upward bias in realized rates of return is synonymous to that in the historical equity premium. This is because the risk-free rate is known at time t, given the subjective disaster risk p_t. Therefore, the difference between the realized and expected rates of return is equal to the difference between the historical and expected equity premium.

those in the data for the US. Besides showing that the model matches the level of returns, this figure shows that the model matches the fluctuations. The model accounts for high realized returns in the 1950s and 1960s through a steady decline in risk during that period. Risk notably ticks up in the 1970s, due to significant market crashes in developing markets (e.g. Argentina, see Figure 3) and the crash due to labor unrest in the U.K. Indeed, this was part of the concern that destabilized markets over this period, and led to a long period of surprisingly low ex post realized returns in the U.S., as the investors adjusted to greater crash risk. Again, in the late 1980s and early 1990s, the U.S. began to markedly diverge from the rest of the world as, within the context of our model, investors concluded that the U.S. was special. In 2008, crash risk again rose, but in tandem across countries; the U.S. did not "catch up." The model explains the market's recovery from the crisis, as well as the shock in 2020. As noted earlier, overall, the correlation between model-implied realized returns and realized-returns in the data across countries is 50%.

The model implied returns match the realized returns in levels. Note that this match is a result of the model, not by construction. Specifically, the rare disaster model falls several percentage points short of the true equity premium. This difference is almost exactly made up by the luck and learning components. A similar result holds across countries– Figure 8 shows that the absolute value of the difference between the predicted and realized mean return clusters are close to zero. Moreover, fluctuations in realized returns match those of the data. With the inferred disaster risk as the only state variable, the realized returns in the model show less variation than those in the data, nevertheless they comove at a business-cycle frequency. Overall, our model is able to capture quite well the variation in realized returns despite of its parsimony.

To conclude, we note that the return news in the U.S. has been predominately positive. The secular decline in perceived disaster risk as opposed to a stable global disaster risk implies that the historical U.S. equity premium is in part due to positive return surprises rather than compensation for risk. Our analysis suggests that the U.S. equity premium puzzle, defined as the difference between the average stock return over and above a short-duration fixed income investment, has been overstated by one-third.

VI. Conclusion

Motivated by the finding in Goetzmann and Jorion (1999) that U.S. equity performance appears to be an exception rather than the norm, we investigate the size of the equity premium puzzle in a comprehensive data set on international equity returns. In particular, we evaluate the problem from the perspective of an investor in 1920 at which time the U.S. equity market had not yet developed into the large and successful market that it is today. We argue that measuring the equity premium based on realized U.S. stock market returns leads to an overestimation of the ex ante expected return by about 1/3 or 200 basis points.

To this end, we model the subjective crash belief of an investor who infers the crash risk in the U.S. by cross learning from other countries. In our framework, the subjective crash risk in the US, after correcting for survivorship bias, is strictly higher than the historical frequency of such events. Moreover, the U.S. experience gradually deviates from that of the global average.

We discuss the asset pricing implication of our findings in an equilibrium setting. First, the secular decline in perceived disaster risk implies a lower equity premium going forward, resulting in a slow-moving upward trend in the observed valuation ratios. Second, the benefit of global diversification is reduced as global economies have become increasingly integrated. As such, downside co-movement is a potentially important risk to manage for investors going forward. Lastly, the U.S. equity premium puzzle should in fact be a smaller puzzle, as the historical equity premium reflects that U.S. investors are surprised on average by the favorable dividend realizations, and the realized excess return results from positive return news rather than compensation for risk. Future finance research should take the bias into account in formulating asset pricing models.

REFERENCES

Avdis, Efstathios, and Jessica A. Wachter, 2017, Maximum likelihood estimation of the equity premium, *Journal of Financial Economics* 125, 589–609.

Backus, David, Mikhail Chernov, and Ian Martin, 2011, Disasters Implied by Equity Index Options, *Journal of Finance* 66, 1969–2012.

Baley, Isaac, and Laura Veldkamp, 2021, Bayesian learning, Working Paper 29338, National Bureau of Economic Research.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823–866.

Betancourt, Michael, Simon Byrne, Sam Livingstone, and Mark Girolami, 2017, The geometric foundations of Hamiltonian Monte Carlo, *Bernoulli* 23, 2257–2298.

Bialkowski, Jedrzej Pawel, and Ehud I. Ronn, 2016, Financial Markets in the Face of the Apocalypse, Working Papers in Economics 16/14, University of Canterbury, Department of Economics and Finance.

van Binsbergen, Jules H., 2020, Duration-based stock valuation: Reassessing stock market performance and volatility, Working Paper 27367, National Bureau of Economic Research.

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

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Chib, Siddhartha, 1998, Estimation and comparison of multiple change-point models, *Journal of Econometrics* 86, 221–241.

Cochrane, John H., 1997, Where is the market going? Uncertain facts and novel theories, *Economic Perspectives* 21, 3–37.

Cogley, Timothy, and Thomas J. Sargent, 2008, Anticipated utility and rational expectations as approximations of Bayesian decision making, *International Economic Review* 49, 185–221.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer, 2016, Parameter learning in general equilibrium: The asset pricing implications, *American Economic Review* .

Constantinides, George M., 2008, Discussion of "macroeconomic crises since 1870", *Brookings Papers on Economic Activity* no. 1, 336–350.

Dew-Becker, Ian, Stefano Giglio, and Bryan Kelly, 2021, Hedging macroeconomic and financial uncertainty and volatility, *Journal of Financial Economics* 142, 23–45.

Dimson, Elroy, Paul Marsh, and Mike Staunton, 2002, Exchange rates and common-currency returns, in *Triumph of the Optimists*, 101 Years of Global Investment Returns, 91–104 (Princeton University Press).

Farmer, Leland E., Emi Nakamura, and Jon Steinsson, 2023, Learning about the long run, forthcoming *Journal of Political Economy.*

Gabaix, Xavier, 2012, Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance, *Quarterly Journal of Economics* 127, 645–700.

Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin, 2013, *Bayesian data analysis, third edition*.

Gelman, Andrew, and Donald B. Rubin, 1992, Inference from Iterative Simulation Using Multiple Sequences, *Statistical Science* 7, 457–472.

Goetzmann, William N, and Philippe Jorion, 1999, Global Stock Markets in the Twentieth Century, *Journal of Finance* 54, 953–9532.

Hamilton, J D, 1994, *Time Series Analysis* (Princeton University Press).

Hansen, Lars Peter, and Kenneth J. Singleton, 1982, Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models, *Econometrica* 50, 1269.

Hoffman, Matthew D., and Andrew Gelman, 2011, The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, *Journal of Machine Learning Research* 15, 1593–1623.

Jones, Christopher S., and Jay Shanken, 2005, Mutual fund performance with learning across funds, *Journal of Financial Economics* 78, 507–552.

Julliard, Christian, and Anisha Ghosh, 2012, Can rare events explain the equity premium puzzle?, *The Review of Financial Studies* 25, 3037–3076.

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran, 2020, The tail that wags the economy: Beliefs and persistent stagnation, *Journal of Political Economy* 128, 2839–2879.

Kreps, David M., 1997, Anticipated Utility and Dynamic Choice (1997), in *Frontiers of Research in Economic Theory*, 242–274 (Cambridge University Press).

Lettau, Martin, Sydney C Ludvigson, and Jessica A Wachter, 2008, The declining equity premium: What role does macroeconomic risk play?, *Review of Financial Studies* 21, 1653–1687.

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Nakamura, Emi, Jón Steinsson, Robert Barro, and José Ursúa, 2013, Crises and recoveries in an empirical model of consumption disasters, *American Economic Journal: Macroeconomics* 5, 35–74.

Neal, Radford M., 2011, MCMC using hamiltonian dynamics, *Handbook of Markov Chain Monte Carlo* 113–162.

Pástor, Ľuboš, and Robert F. Stambaugh, 2001, The equity premium and structural breaks, *Journal of Finance* 56, 1207–1239.

Pastor, Lubos, and Pietro Veronesi, 2009, Learning in financial markets, *Annual Review of Financial Economics* 1, 361–381.

Rietz, Thomas A., 1988, The equity risk premium a solution, *Journal of Monetary Economics* 22, 117–131.

Stambaugh, Robert F., 2011, Inference about survivors, *Quarterly Journal of Finance* 1, 423–464.

Tsai, Jerry, and Jessica A. Wachter, 2015, Disaster Risk and Its Implications for Asset Pricing, *Annual Review of Financial Economics* 7, 219–252.

Wachter, Jessica A., 2013, Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?, *Journal of Finance* 68, 987–1035.

Wachter, Jessica A., and Yicheng Zhu, 2023, Learning with rare disasters, Working paper, University of Pennsylvania.

Weitzman, Martin L., 2007, Subjective expectations and asset-return puzzles, *American Economic Review* 97, 1102–1130.

Welch, Ivo, 2016, The (time-varying) importance of disaster risk, *Financial Analysts Journal* 72, 14–30.

Figure 1. Risk Return Trade-off

The upper panel plots the arithmetic return against volatility while the lower panel plots the geometric return against volatility (all adjusted for unequal sample lengths except for geometric return). The geometric (buy-and-hold) return is the annualized rate of return one would earn by investing $1 when the market comes into existence and holding the position till the end of the sample. For the few countries which experienced breaks in its stock market, we assume that $1 more is invested at the point of restart. The median regression line is shown in blue.

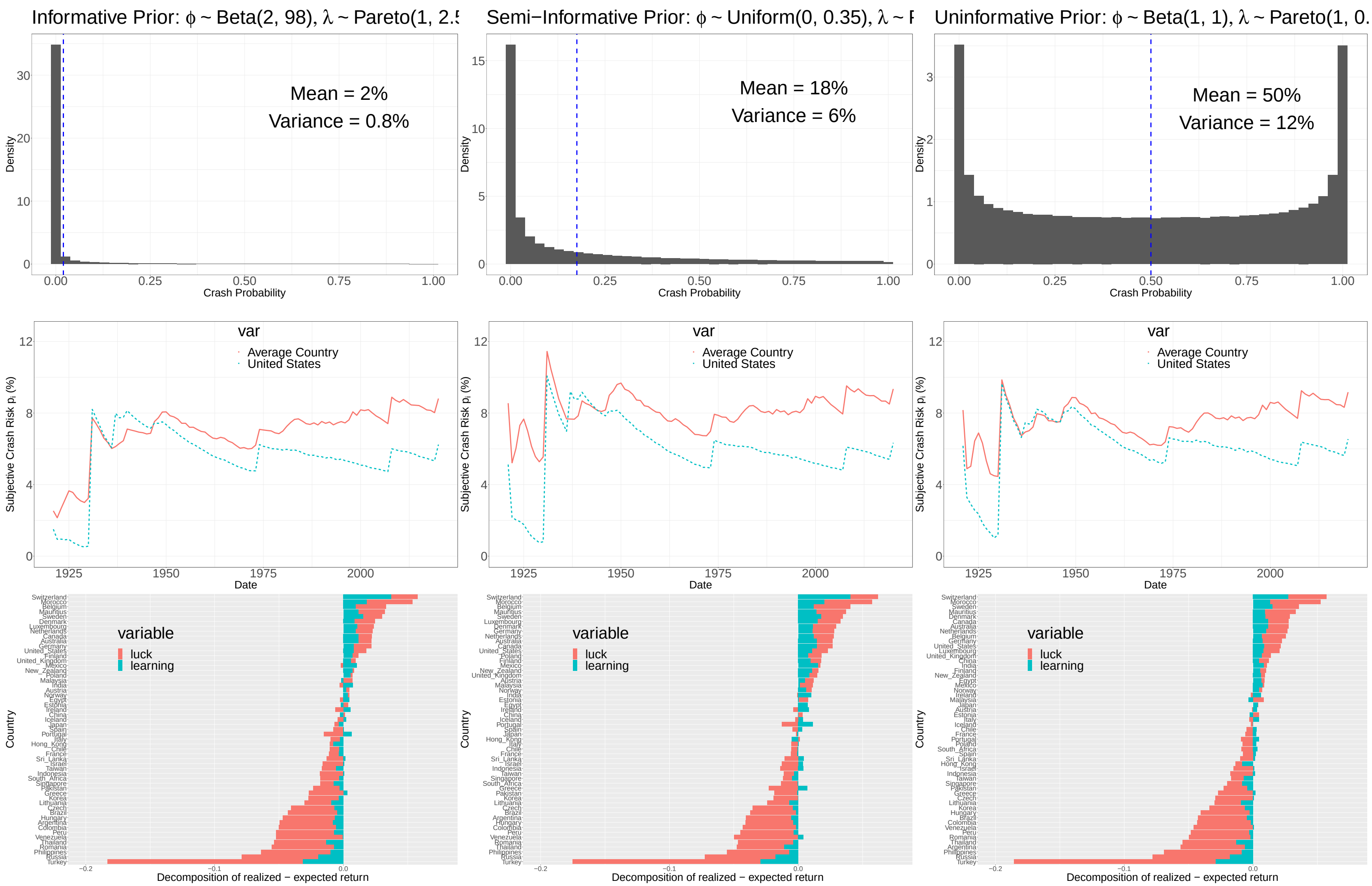

Figure 2. Robustness to Different Prior Specifications

Each column of the figure corresponds to one specification of the hyperprior: *informative, semi-informative* and *uninformative*. The first row plots the prior distribution of idiosyncratic crash risk implied by the specific hyperprior distribution. Gray bars denote the sampling distribution of 1,000,000 draws, and blue dotted lines denote the mean. The second row depicts the evolution of subjective crash risk under each hyperprior specification following the model in II.A. The bottom row shows the asset pricing results derived according to the model in V.

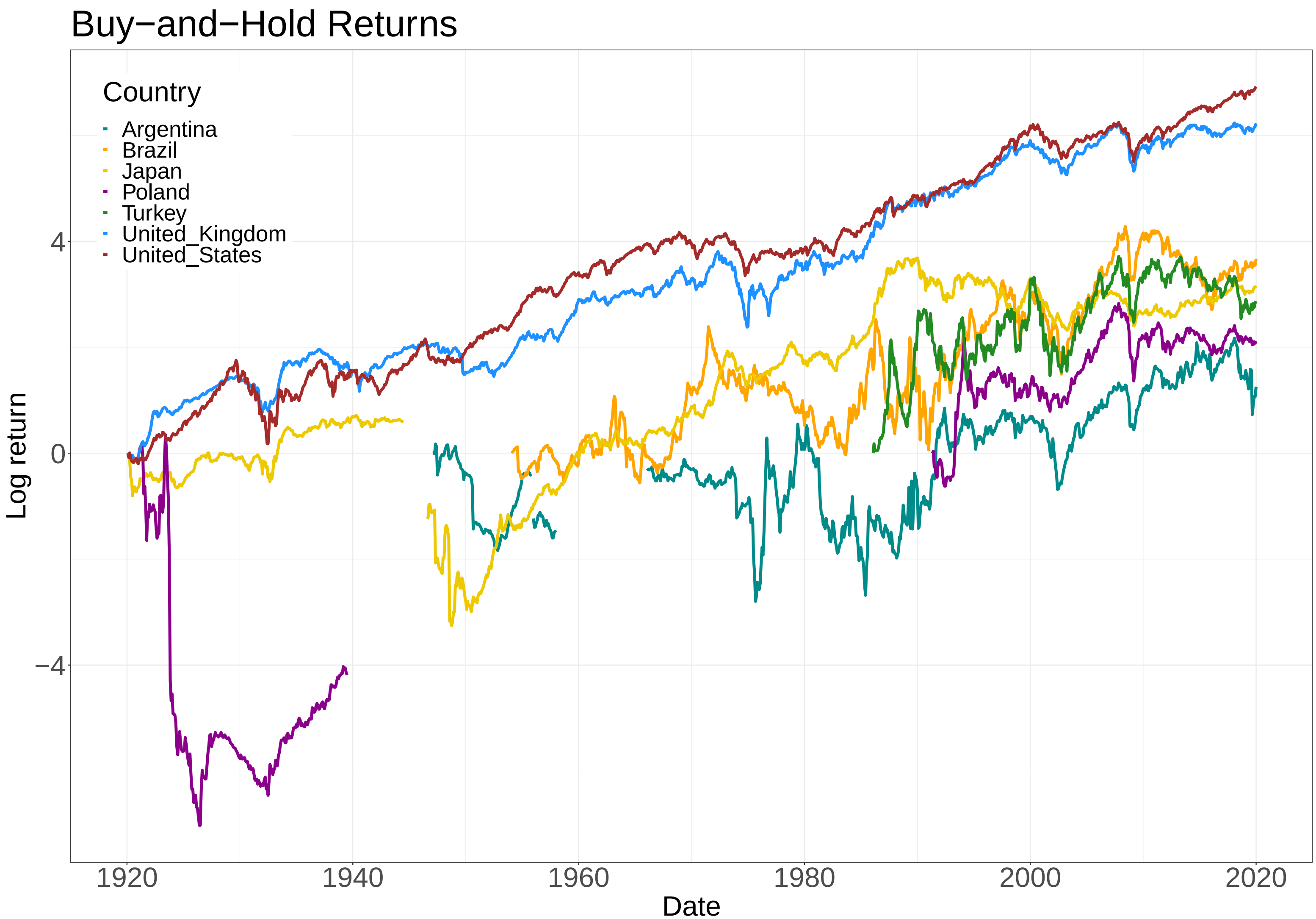

Figure 3. Buy-and-Hold Return

The figure plots the buy-and-hold return of investing \$1 in max{1920, when the market enters the sample} for selected countries. For the few countries (e.g. Poland) which had its stock market nationalized, it is assumed that \$1 more is invested at the point of restart. The U.S. stands out in having a smooth and continuous series, whereas other markets experience trading breaks and high volatility.

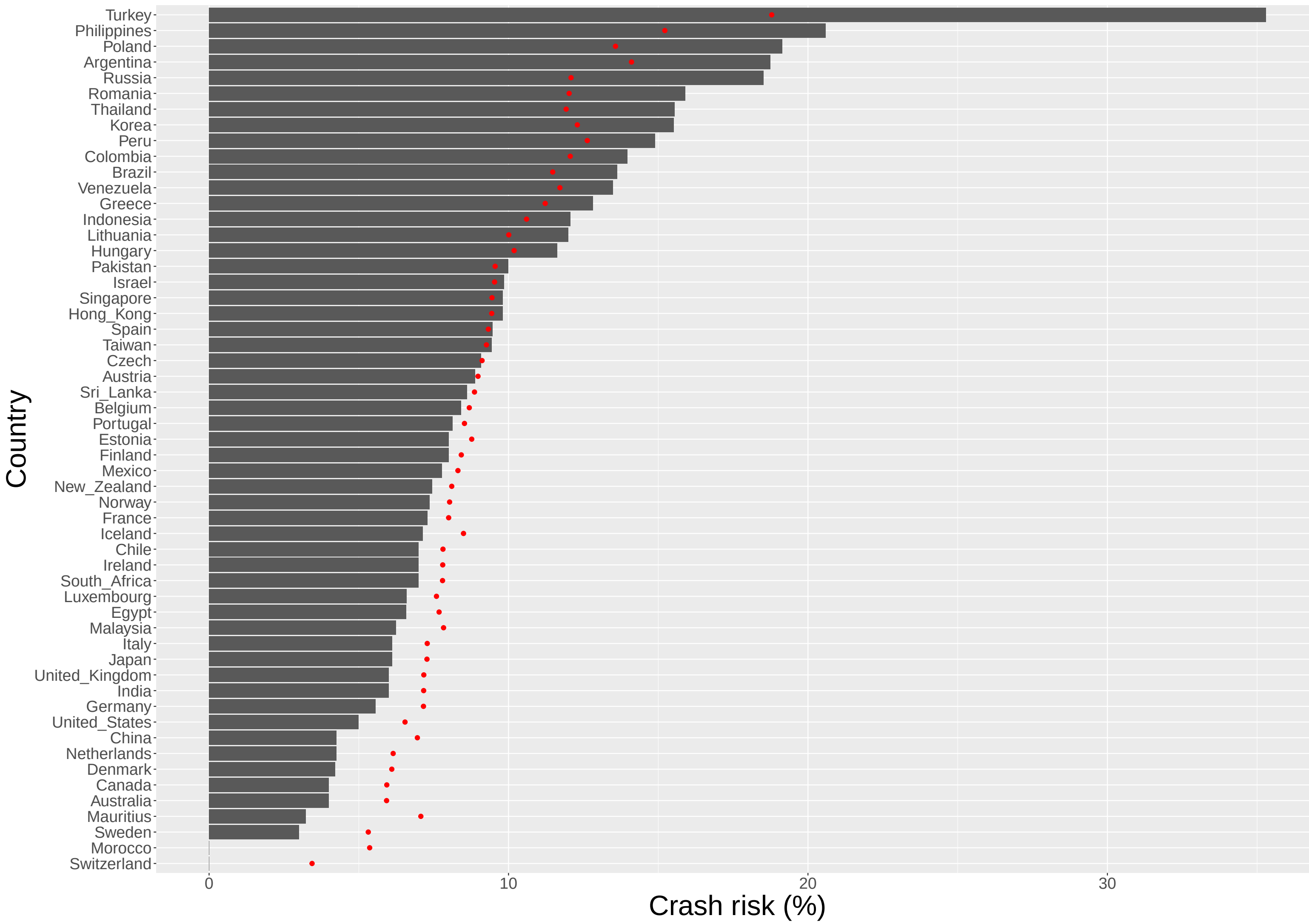

Figure 4. Subjective Crash Risk v.s. In-sample Frequency

The figure plots the posterior mean crash risk (%) at the end of the sample (March 2020) against the historical frequency (%) of market crashes for each country. The posterior estimates are derived from the model in Section II.A. The red dot denotes the posterior mean crash risk estimate. The gray bar illustrates the historical frequency of market crashes in the sample.

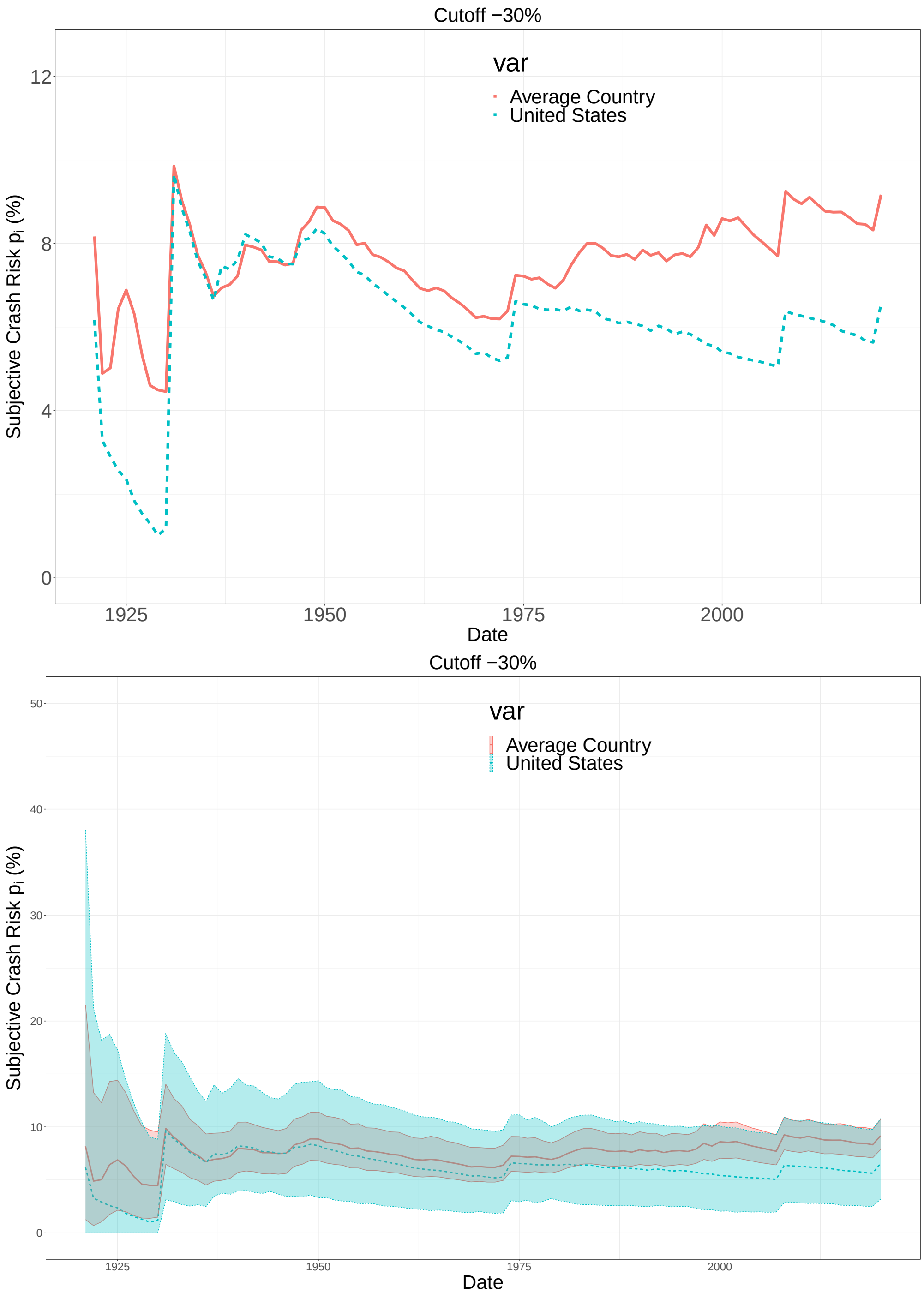

Figure 5. Subjective Crash Risk – 1921-2020

The figure shows the time series evolution of posterior mean crash risk over the full sample from the model in Section II.A. The upper panel plots the posterior mean while the lower panel also provides the 95% simulation-based credible set. The implied global mean crash risk $\alpha/(\alpha+\beta)$ is denoted by the solid red line. The U.S. crash risk is shown in the dashed blue line.

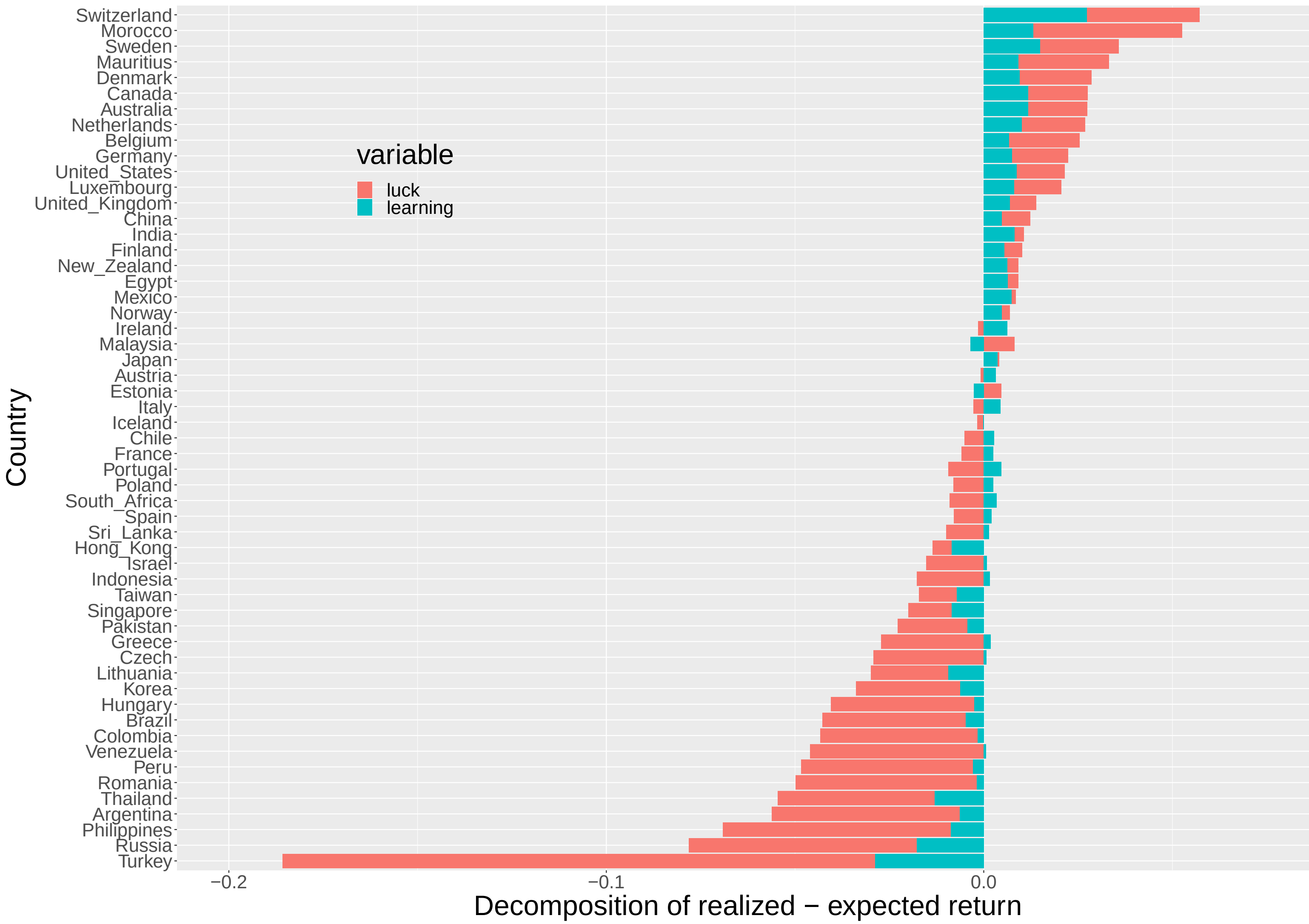

Figure 6. Decomposition of the Equity Premium

The figure decomposes the equity premium in each country into lucky realizations of dividend growth (*luck*) and learning induced positive return news (*learning*). We calibrate the model in Section V with $\gamma = 3$, $\beta = 0.98$, $\mu = 0.0252$, $\sigma = 0.02$, $\lambda = 2.6$, $\psi = 1$, and $b = 0.3$, while feeding in the posterior crash belief over time (from 1931 to 2020) as the perceived disaster risk. We compare the average realized returns with the expected returns.

Figure 7. 10-year Rolling Average Returns in the US

The figure plots the 10-year rolling average returns in the U.S. stock market based on the model (red solid) and in the data (blue dashed). We calibrate the model in Section V with $\gamma = 3$, $\beta = 0.98$, $\mu = 0.0252$, $\sigma = 0.02$, $\lambda = 2.6$, $\psi = 1$, and $b = 0.3$, while feeding in the posterior crash belief over time (from 1931 to 2020) as the perceived disaster risk. We compare the average realized returns with the expected returns.

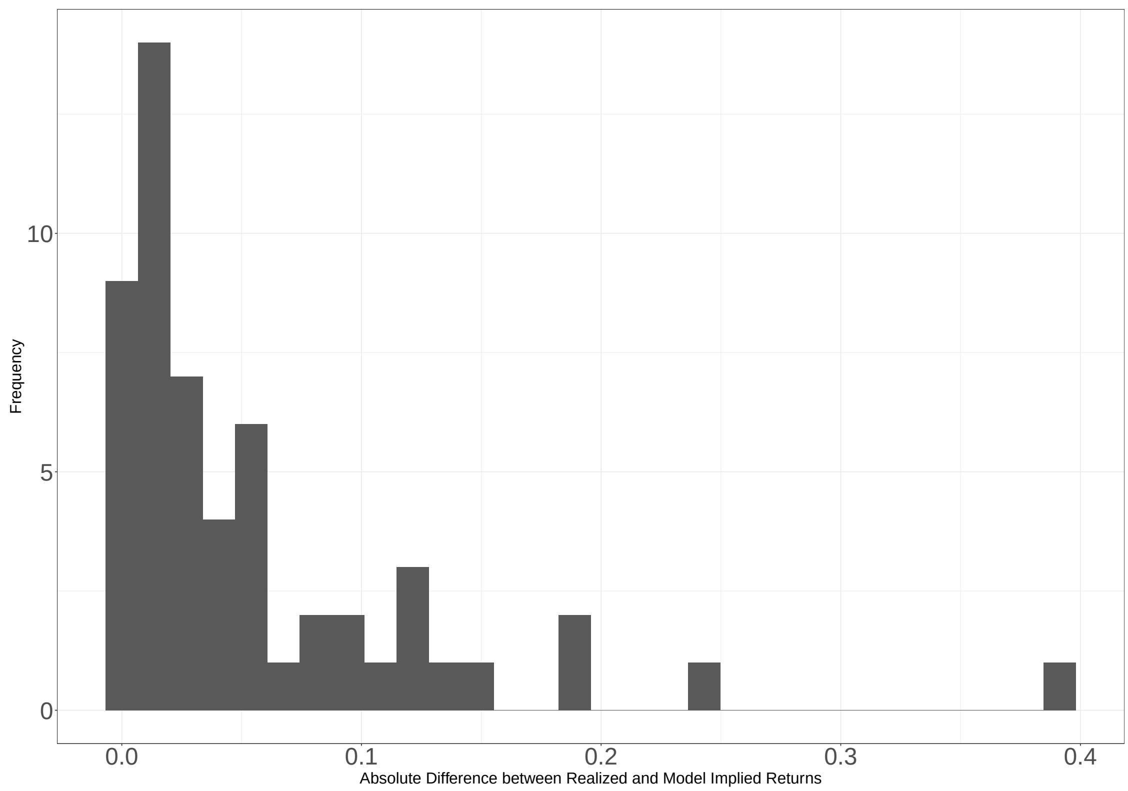

Figure 8. Absolute difference between predicted and realized mean return

The figure shows the absolute difference between the average return predicted by the model and in the data across countries. We calibrate the model in Section V with $\gamma = 3$, $\beta = 0.98$, $\mu = 0.0252$, $\sigma = 0.02$, $\lambda = 2.6$, $\psi = 1$, and $b = 0.3$, while feeding in the posterior crash belief over time (from 1931 to 2020) as the perceived disaster risk. We compare the average realized returns with the expected returns

Table I: Summary Statistics – Global Equity Returns

The table presents the summary statistics (mean arithmetic return, standard deviation, maximum, minimum, skewness, kurtosis, geometric return, beta with respect to an equal weighed global portfolio, and beta with respect to a capitalization weighed global portfolio) for annual stock returns across 55 markets from 1920 to March 23, 2020. Returns are measured as changes in total return indices from the year end of the previous year to the current year. The total return indices are constructed assuming that dividends are reinvested monthly. Mean arithmetic return and standard deviation are adjusted for unequal sample lengths.

Country	R^A (%)	**Std (%)**	**Max (%)**	**Min (%)**	**Skewness**	**Kurtosis**	R^G (%)	β^{equal}	β^{cap}
Argentina	21.62	115.43	789.11	-80.64	5	32.52	0.98	1.71	1.48
Australia	9.59	23.83	92.2	-51.89	0.32	4.02	6.94	0.85	1
Austria	9.6	46.31	259.83	-100	2.34	13.41	2.97	1.07	0.79
Belgium	4.5	25.86	71.55	-100	-0.32	5.39	3.23	0.84	0.97
Brazil	19.36	61.5	277.07	-74.3	1.39	6.29	4.36	1.38	1.23
Canada	7.95	20.99	71.26	-45.18	-0.02	3.23	5.67	0.73	0.96
Chile	10.59	36.13	130.39	-52.75	1.05	4.29	4.9	1.06	0.88
China	9.22	40.01	135.09	-100	0.68	4.99	1.66	0.89	1
Colombia	5.19	37.82	196.05	-55.9	1.99	9.83	-0.22	0.78	0.49
Czech	13.82	40.96	119.67	-100	0.29	3.82	2.62	1.1	0.79
Denmark	7.54	25.79	105.44	-100	0.05	7.4	5.97	0.81	0.85
Egypt	10.74	38.55	157.06	-53.8	1.52	6.06	2.63	1.19	1.17
Estonia	20.69	54.31	216.12	-64.62	1.65	7.93	11.32	1.14	0.91
Finland	11.47	33.19	123.91	-53.48	0.98	4.43	6.5	0.94	0.99
France	5.74	29.14	77.88	-100	-0.07	4.23	3.86	0.94	1.06
Germany	13.41	60.24	428.3	-100	4.4	31.09	6.12	0.77	0.75
Greece	15.19	57.81	269.56	-83.56	2.12	9.28	3.68	1.56	1.51
Hong Kong	16.69	41.2	149.96	-61.25	0.78	4.2	9.75	1.08	1.36
Hungary	7.78	39.2	121.17	-100	0.16	4.05	1.23	1.02	1.03
Iceland	12.62	38.64	110.56	-94.73	-0.17	4.38	0.66	0.77	1
India	9.84	32.9	166.56	-60.54	1.75	8.8	5.07	1.01	1.09
Indonesia	9.93	49.82	241.78	-73.89	1.95	9.67	-0.09	1.03	1.06

Country	R^A (%)	Std (%)	Max (%)	Min (%)	Skewness	Kurtosis	R^G (%)	β^{equal}	β^{cap}
Ireland	7.17	28.09	89.89	-56.04	0.61	3.76	3.36	0.9	1.05
Israel	6.55	33.66	88.85	-71.64	0.47	3.25	0.51	0.82	0.9
Italy	7.31	32.42	127.83	-100	0.61	5.9	2.53	0.87	1
Japan	8.63	33.02	140.23	-100	0.68	6.28	2.88	0.71	0.84
Korea	16.32	43.48	137.82	-67.28	0.71	3.52	11.03	0.79	1.04
Lithuania	13.09	40.17	142.17	-65.75	1.07	5.82	6.45	1.18	1.11
Luxembourg	8.21	27.95	122.39	-59.97	0.75	5.61	4.65	0.98	1.01
Malaysia	6.78	30.44	104.51	-67.7	0.4	4.1	2.62	0.76	0.81
Mauritius	11.24	30.89	84.99	-43.43	0.52	2.53	7.52	0.75	0.71
Mexico	10.53	42.29	238.54	-88.45	2.2	11.8	3.13	1.08	0.96
Morocco	9.73	23.25	71.82	-25.64	0.58	2.69	7.57	0.48	0.46
Netherlands	7.33	24.48	78.58	-100	-0.52	6.69	5.26	0.82	0.99
New Zealand	8.29	28.78	141.64	-49.67	1.71	9	5	0.86	0.85
Norway	6.87	29.02	96.69	-100	0	5.12	4.17	0.91	0.85
Pakistan	11.67	39	167.83	-65.89	1.43	7.24	5.54	0.75	0.58
Peru	15.23	66.78	432.75	-76.95	3.29	18.8	2.82	1.69	1.14
Philippines	6.57	46.88	201.59	-71.38	1.33	6.42	-2.79	1.4	1.31
Poland	19.84	121.57	780.01	-100	5.31	34.04	1.62	2.32	1.71
Portugal	10.85	41.97	172.47	-90.84	1.52	7.01	1.78	1	1.09
Romania	11.84	63.05	225.5	-100	1.39	6.01	0.7	1.57	1.26
Russia	32.42	68.63	188.48	-84.81	0.45	2.58	13.56	1.62	1.58
Singapore	11.75	42.5	207.41	-53.65	1.98	10.07	5.3	1.11	1.37
South Africa	9.72	28.76	155.15	-47.33	1.38	8.41	6	0.97	0.96
Spain	7.74	29.96	151.33	-46.98	1.17	7.29	3.12	0.9	0.93
Sri Lanka	10.62	33.8	124.99	-38.78	1.51	5.74	3.01	0.6	0.34
Sweden	9.47	23.68	70.74	-49.1	0.15	3.26	6.51	0.82	0.98
Switzerland	8.95	21.31	92.76	-29.96	0.78	4.74	6.47	0.66	0.81
Taiwan	15.54	44.79	172.36	-64.44	1.15	4.94	7.83	1.09	1.13
Thailand	15.19	45.54	140.16	-76.37	0.6	3.43	5.44	1.16	1.23
Turkey	36.03	116.38	515.55	-62.44	2.42	9.78	7.5	2.73	2.73

Country	R^A (%)	**Std (%)**	**Max (%)**	**Min (%)**	**Skewness**	**Kurtosis**	R^G (%)	β^{equal}	β^{cap}
United Kingdom	8.85	25.36	103.36	-56.6	0.77	5.76	5.82	0.79	1.05
United States	8.96	20	56.78	-38.62	-0.33	2.83	6.76	0.56	0.94
Venezuela	15.72	64.18	446.89	-68.79	3.9	24.81	5.5	0.53	0.12

Table II: Summary Statistics – Frequency Of Market Crashes

The table presents the periods in which each market is in the sample. A market is deemed to exist if it is present in our sample at the end of the year. n^C *is the number of market crashes worse than* -30% *that took place when the market is in the sample.* N *is the total number of periods a market is in the sample.* p *is the in-sample frequency of market crashes.*

Country	Coverage	n^C	N	$p(\%)$	Country	Coverage	n^C	N	$p(\%)$
Argentina	1948-1957, 1967-2020	12	64	18.75	Luxembourg	1930-2020	6	91	6.59
Australia	1920-2020	4	100	4	Malaysia	1973-2020	3	48	6.25
Austria	1923-1938, 1947-2020	8	90	8.89	Mauritius	1990-2020	1	31	3.23
Belgium	1920-1940, 1946-2020	8	95	8.42	Mexico	1931-2020	7	90	7.78
Brazil	1955-2020	9	66	13.64	Morocco	1980-2020	0	41	0
Canada	1920-2020	4	100	4	Netherlands	1920-1940, 1947-2020	4	94	4.26
Chile	1920-2020	7	100	7	New Zealand	1927-2020	7	94	7.45
China	1920-1941, 1995-2020	2	47	4.26	Norway	1920-1940, 1946-2020	7	95	7.37
Colombia	1928-2020	13	93	13.98	Pakistan	1961-2020	6	60	10
Czech	1920-1938, 1995-2020	4	44	9.09	Peru	1927-2020	14	94	14.89
Denmark	1920-1940, 1946-2020	4	95	4.21	Philippines	1953-2020	14	68	20.59
Egypt	1920-1968, 1993-2020	5	76	6.58	Poland	1922-1939, 1992-2020	9	47	19.15
Estonia	1996-2020	2	25	8	Portugal	1932-1974, 1978-2020	7	86	8.14
Finland	1920-2020	8	100	8	Romania	1927-1947, 1998-2020	7	44	15.91
France	1920-1940, 1945-2020	7	96	7.29	Russia	1994-2020	5	27	18.52
Germany	1920-1933, 1958-2020	6	76	7.89	Singapore	1970-2020	5	51	9.8
Greece	1929-1931, 1933-1939, 1953-2020	10	78	12.82	South Africa	1920-2020	7	100	7
Hong Kong	1970-2020	5	51	9.8	Spain	1920-1935, 1941-2020	9	95	9.47
Hungary	1925-1930, 1933-1940, 1992-2020	5	43	11.63	Sri Lanka	1953-1974, 1985-2020	5	58	8.62
Iceland	1993-2020	2	28	7.14	Sweden	1920-2020	3	100	3
India	1920-2020	6	100	6	Switzerland	1920-2020	0	100	0
Indonesia	1925-1939, 1978-2020	7	58	12.07	Taiwan	1968-2020	5	53	9.43
Ireland	1920-2020	7	100	7	Thailand	1976-2020	7	45	15.56
Israel	1950-2020	7	71	9.86	Turkey	1987-2020	12	34	35.29
Italy	1920-1943, 1946-2020	6	98	6.12	United Kingdom	1920-2020	6	100	6
Japan	1920-1944, 1947-2020	6	98	6.12	United States	1920-2020	5	100	5
Korea	1963-2020	9	58	15.52	Venezuela	1930-1940, 1943-2020	11	89	12.36
Lithuania	1996-2020	3	25	12					

Table III: Data Source – Global Financial Data

The table provides the tickers used to retrieve data from the Global Financial Data(GFD) database. The total return series are the main data series we used. When not available, a representative total return index is constructed by combining price level and dividend yield. We fill in missing values of dividend yield with the sample average and interpolate annual dividend yield into monthly dividend yield. The representative total return series are constructed assuming monthly reinvestment of dividend. We replace the total return index for the U.S. with that from CRSP because CRSP accounts for the universe of U.S. stocks while GFD proxies it with S&P500.

Country	Price Level	Dividend Yield	Total Return	Country	Price Level	Dividend Yield	Total Return
Argentina	GFPRARGSTD	SYARGYM		Luxembourg	GFPRLUXSTD	SYLUXYM	GFTRLUXSTD
Australia			GFTRAUSSTD	Malaysia			GFTRMYSSTD
Austria	GFPRAUTSTD	SYAUTYM		Mauritius			GFTRMUSSTD
Belgium			GFTRBELSTD	Mexico	GFPRMEXSTD	SYMEXYM	
Brazil			GFTRBRASTD	Morocco	GFPRMARSTD	SYMARYM	
Canada			GFTRCANSTD	Netherlands	GFPRNLDSTD	SYNLDYAM	
Chile	GFPRCHLSTD	SYCHLYM		New Zealand	GFPRNZLSTD	SYNZLYM	
China	CNSHFVUSDA	SYCHNYM	GFTRCHNSTD	Norway	GFPRNORSTD	SYNORYM	
Colombia	GFPRCOLSTD	SYCOLYM		Pakistan	GFPRPAKSTD	SYPAKYM	
Czech	GFPRCZESTM, GFPRCZESTD	SYCZEYM		Peru	GFPRPERSTD	SYPERYM	
Denmark			GFTRDNKSTD	Philippines	GFPRPHLSTD	SYPHLYM	
Egypt	GFPREGYSTM, GFPREGYSTD	SYEGYYM		Poland	GFPRPOLSTM	SYPOLYM	GFTRPOLSTD
Estonia			GFTRESTSTD	Portugal	GFPRPRTSTD	SYPRTYM	
Finland			GFTRFINSTD	Romania	GFPRROUSTD, ROBUCHM	SYROUYM	
France			GFTRFRASTD	Russia	GFPRRUSSTD	SYRUSYM	
Germany			GFTRDEUSTD	Singapore			GFTRSGPSTD
Greece	GFPRGRCSTD, GRATHENM	SYGRCYM		South Africa	GFPRZAFSTD	SYZAFYM	
Hong Kong			GFTRHKGSTD	Spain	GFPRESPSTD	SYESPYM	
Hungary	GFPRHUNSTM	SYHUNYM	GFTRHUNSTD	Sri Lanka	GFPRLKASTD, GFPRLKASTM	SYLKAYM	
Iceland	GFPRISLSTD	SYISLYM	GFTRISLSTD	Sweden			GFTRSWESTD
India			GFTSINDM	Switzerland	GFPRCHESTD	SYCHEYM	
Indonesia	GFPRIDNSTD, GFPRIDNSTM	SYIDNYM	GFTRIDNSTD	Taiwan	GFPRTWNSTD	SYTWNYM	
Ireland			GFTSIRLR	Thailand			GFTRTHASTD
Israel			GFTSISRM	Turkey			GFTRTURSTD
Italy	GFPRITASTD	SYITAYM		United Kingdom			GFTRGBRSTD
Japan			GFTRJPNSTD	United States			CRSP
Korea			GFTRKORSTD	Venezuela	GFPRVENSTD	SYVENYM	
Lithuania			GFTRLTUSTD				

Table IV: Decomposition of the U.S. Equity Premium

The table presents the decomposition of average returns into ex ante expected returns, luck and learning as in Equation 8, in percent. We repeat the analysis for the three different prior specifications we consider. The first row corresponds to the informative prior, the second to the semi-informative prior and the last one to the uninformative prior.

	Average Realized Return	Ex Ante Expected Return + Residual	Luck	Learning
Informative Prior	8.96	7.19	0.96	0.81
Semi-informative Prior	8.96	6.67	1.21	1.08
Uninformative Prior	8.96	6.76	1.23	0.97