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

Jules van Binsbergen[†] Sophia Hua[‡] Jessica A. Wachter[§]

October 2020

ABSTRACT

We quantify the extent of survivorship bias in the US equity market performance and find that it explains about 1/3 of the equity risk premium in the past century. The wedge between the realized and expected excess return can be attributed to *luck* and *learning*. We model the subjective crash belief of an investor who infers the crash risk in the US by cross learning from other countries. The investor distinguishes between idiosyncratic and systematic crash risk and identifies a structural break in the magnitude of the systematic risk at the turn of the century. The crash belief in the US shows a persistent and widening divergence from the implied global average. Consistent with the model, we also document that a global CAPM fits well the cross-section of country average returns.

*We thank Sid Chib, Winston Dou, Alice Gindin, Karen Lewis, Nick Roussanov, Frank Schorfheide, Robert Stambaugh, and seminar participants at NBER-NSF SBIES for comments.

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

[‡]The Wharton School, University of Pennsylvania; huaye@wharton.upenn.edu.

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

I. Introduction

The 6% average annual outperformance of stocks over and above the returns on a short duration risk-free investment has attracted ample attention in the economics and finance literature. Because this outperformance was initially hard to explain using prevailing consumption-based asset pricing models, it is now known as the equity premium puzzle (Mehra and Prescott (1985) and Hansen and Singleton (1982)). Since its introduction, many different potential solutions have been proposed.[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 statistical robustness.[2] This paper contributes to this latter approach and builds on the striking finding in Goetzmann and Jorion (1999) that US 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 driving the outperformance of markets still in existence. In this case, looking only at the best-performing country yields a biased estimate of the ex ante unconditional return on the equity market.[3] Moreover, 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 US equity market.

In this paper, we construct a comprehensive database with total return indices from a cross-section of 55 countries in the past century (1920-2020) and revisit the question of whether survivorship bias plays an important role in the historical equity premium. While we find that equity has outperformed short-duration fixed income claims across countries

[1] For an incomplete summary, see Campbell and Cochrane (1999), Wachter (2013), Bansal and Yaron (2004) 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 which constitute a substantial fraction of realized returns, particularly in the early part of our sample.

in our sample, we also find that the US appears lucky compared to most other countries. After all, the US was never heavily implicated in either world war. The few crises that took place, including the Cuban Missle Crisis, resolved themselves peacefully without straining the economy. As such, from the perspective of an investor in 1920, it would be impossible to predict with certainty the rise and ensuing dominance of the US equity market. In fact, Argentina was among the world's ten richest countries back then, 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 Argentina stocks at the turn of the century. Even though investors may have expected a high return, they ended up earning a low realized return ex post, despite high levels of volatility.[4] In all, we find that survivorship bias alone accounts for about 1/3 of the measured equity risk premium.

We also document that a global CAPM can explain at least part of the difference in average returns in the cross-section of countries, despite its well-documented failure in the cross-section of average US stock returns. Figure 1 plots the average annual stock return (expressed in real dollar terms) against the exposure to a capitalization weighted global portfolio for every market in our sample. A median regression fitted to the data implies a 7.1% global equity premium and a 3% risk-free rate. In addition, the US lies very close to the security market line (SML). With a beta slightly below 1, its performance does not look exceptional compared to other countries in the sample. Perhaps somewhat surprisingly, the developed stock markets uniformly lie on or slightly below the SML, whereas most emerging markets are above the SML due to high volatility. A standard regression fitted to the data has a positively sloped SML with an R-squared of 50% providing support for a positive risk

[4]When comparing returns across markets, it matters whether one is using geometric returns (GR) or arithmetic returns (AR). By the Cauchy-Schwarz inequality, GR $\leq$ AR. If one assumes log-normality, AR = GR + $1/2\sigma^2$. Figure 2 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.

return relationship in international markets.

To quantify the effect of potential survivorship bias on measured 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 -30%. Without loss of generality, we choose to model learning about the frequency rather than the magnitude of disasters. Our model is easily adaptable for any severity of market crashes. It is possible to estimate the subjective crash belief for a continuum of cutoffs and obtain a subjective return distribution in terms of a kernel density function.[5] 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 from crashes taken place in other countries about crash risk in the United States. 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 in a country's experience (to what extent 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. This is reasonable because, if we have a sufficiently long sample, we would optimally put our entire faith on country-specific observations to determine the frequency of rare events.

We further incorporate systematic and idiosyncratic crash risk to account for the conditional dependence of market crashes observed in the data. The augmented Bayesian model involves an investor learning about the probability of the world switching into a high or low systematic risk regime, the incremental likelihood of any country to experience a stock

[5] An alternative non-parametric approach for modeling returns uses the Dirichlet process. However, for ease of economic interpretation, we focus on one feature of the returns data arguably of most interest to investors: downside crash risk.

market crash in the high systematic risk period, the idiosyncratic crash risk for each country, and the implied global mean crash risk. The model succeeds in pooling information across countries to make inferences regarding systematic crash risk, while cross learning from international crash realizations helps correct for survivorship in the US. Lastly, we embed all the above in a filtering problem, where an investor also infers the latent state of the world, and accounts for potential structural breaks in the magnitude of the systematic risk. We apply a state-of-the-art Markov Chain Monte Carlo (MCMC) algorithm – Hamiltonian Monte Carlo (HMC) – to obtain the posterior crash belief of the investor. We replicate the information set of an investor at each point in time, and track the subjective crash belief of an investor over time.

At the start of our sample we assume an uninformative hyperprior over the global crash risk generating process for all countries. As the sample progresses, the subjective crash belief for the US 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 (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 denote the resulting wedge between the historical and expected risk premium as *learning.* In addition, with a well-defined posterior belief about disaster, we can evaluate the perception of *luck* in the US equity market experience. We show that *luck* and *learning* jointly explain 2% out of the 6% historical equity premium. So indeed, the equity premium puzzle should be a smaller 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. The problem goes hand in hand with the fact that the performance of poor-performing assets, with higher attrition rates, is badly documented. Looking only at the easily available data and hence the surviving sample lead to an overestimation of the 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. Nevertheless, Stambaugh (2011) points out their approach of assuming a complete lack of information and imposing a common prior leads to too strong of a correction to survivorship bias in the mutual fund industry. In the paper, we apply Jones and Shanken (2005)'s approach to discipline the variation of crash risk across countries. The fact that global equity markets were more homogeneous in the 1920s than they are today rationalizes our assumption of a common prior on crash risk. Moreover, the rare-occurring nature of market crashes justifies the need of learning from the cross-section to discipline the estimate of country-specific crash risk. More importantly, it reveals to what extent the luckiness in the US experience can be attenuated by learning across countries.

Our paper is also closely tied to the idea that there are large negative shocks that did not realize in the US sample, while nonetheless the investors are 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 out-of-the-money put options. The over-pricing of the options reveals that the underlying economy is far from Gaussian, but in fact highly negatively skewed. Nevertheless, it is impossible 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 the rare events is also subject to debate (Backus et al. (2011), Welch (2016)). In this regard, our paper provides a well-disciplined benchmark of physical crash belief formation, which allows spillover of crashes taken place in other countries to inform the crash belief in the US.

On the methodological side, the paper combines Bayesian filtering technique with a hierarchical Bayesian model to infer latent states while teasing out country-level heterogeneity. We bring together the literature on regime switch 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, then we discuss the prior choice and derive the posteriors for different sets of parameters of interest. Section III discusses the data and methodology. In Section IV, we interpret the results from the models. Next, we provide the asset pricing implication of the secular decline in crash risk in the past century in Section V. Lastly, we conclude.

II. Model

A. Likelihood

State Transition. The crash risk for country i at time t, $p_{i,t}$, consists of an idiosyncratic component p_i^I, which is different across countries and static, and a systematic part p_t^S, which is shared among all and can be time-varying. The common systematic risk manifests the increased likelihood of all countries to experience a crash in a high risk regime, denoted by $x_t = H$. If $p_t^S = 0$, market crashes are independent across countries. Otherwise, they are cross-sectionally correlated.

$$p_{i,t} = p_i^I + p_t^S \mathbf{1}_{x_t = H} \quad \forall i \tag{1}$$

The economy further switches between high risk and low risk regimes, following the regime switching matrix P. The assumption implies time series correlation, i.e. clustering, in the occurrence of market crashes.

$$P = \begin{bmatrix} p_{HH} & p_{HL} \\ p_{LH} & p_{LL} \end{bmatrix} \tag{2}$$

We also allow a possible structural break to take place in the magnitude of the systematic risk p_t^S. The break occurs at time t_b, and p_t^S takes the values p_1^S and p_2^S pre and post the break respectively.

$$p_t^S = \begin{cases} p_1^S & \text{if } t < t_b \\ p_2^S & \text{if } t \geq t_b \end{cases} \tag{3}$$

Measurement Equation. The investor observes a panel of crash indicators $D_{i,t}$, which realize independently conditionally on $p_{i,t}$.

$$D_{i,t}|p_{i,t} \overset{i.i.d.}{\sim} \text{Bernoulli}(p_{i,t}) \quad \forall i, t \tag{4}$$

Overall, the model captures correlation across countries and time in a parsimonious manner. There are various ways to simply the inference task of the investor. For instance, a naive investor may consider market crashes to be independent across countries by setting $p_t^S = 0$. Alternative, one may abstract away from learning about the latent states by including the state indicators in the information set of the investor.

B. Prior

Equipped with little information about the performance of global equity markets at the start of the sample, the investor in the model has uninformative priors over all parameters. Most of the parameters of interest are probabilities with support between 0 and 1. Therefore, we impose uniform priors on such parameters. The assumption also enables the investor to learn quickly in face of data. We summarize the prior choices in Table III.

In particular, we adopt the conservative assumption that the magnitude of the systematic risk is drawn from a uniform distribution independently pre and post the structural break. Absent a structural break in the data, p_1^S and p_2^S would be indistinguishable from one another, rendering the model a bad fit of the data. The prior belief on the location of the structural break t_b is also uniform from time 1 to τ, as an investor has no advance knowledge when, if at all, a structural break would take place, before observing the data.

Importantly, the correction for survivorship bias comes in through a hierarchical prior on the degree of idiosyncratic crash risk in each country. That is, we assume that the idiosyncratic crash risk p_i are drawn from the same distribution, which is unknown to the investor. The assumption implies that the crash risk across countries are weakly positively correlated. Meanwhile, we do not stipulate the extent of the correlation. The investor instead infers it from the data. Furthermore, such correlation is taken into account by the investor in forming beliefs about idiosyncratic crash risk. We show that some interesting cross-learning effects arise as a result. The hierarchical prior on the idiosyncratic crash risk is as follows.

Prior. We assume that the country-specific crash risk p_i^I is drawn from the same distribution:

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

Hyperprior. The investor has a hyperprior over the parameters α and β, that govern the global crash risk distribution. The uninformative hyperprior is taken from Gelman et al. (2013).[6] The model assumptions above allow for investors to learn about crash risk across countries.

$$f(\alpha, \beta) \propto (\alpha + \beta)^{-5/2}$$

Note that the hierarchical prior assumption introduces two additional parameters, α and β into the model. We denote $\boldsymbol{\theta}$ as the collection of parameters of interest: $\boldsymbol{\theta} = \{\{p_i^I\}_{i=1}^n, p_{LH}, p_{HL}, \alpha, \beta, p_1^S, p_2^S, p_0^H, t_b\}$, where p_0^H is the probability of starting off in a high

[6] We provide further intuition in Appendix A.

systematic risk state.

C. Posterior

For maximum clarity, we separately discuss different ingredients of the models and how the investor infers the relevant parameters. We start with the simple case where markets crashes are conditionally independent across countries. The simplification allows us to explain the cross-learning effect without systematic risk complicating the problem. Building upon the model, we then introduce systematic risk to account for cross-country correlation, while assuming that investor is able to observe the latent states of the world. Lastly, we include a structural break and have the investor filter out the latent states of the world.

C.1. Correction For Survivorship Bias

Without systematic risk, crash realizations $D_{i,t}$ are independently and identically distributed across countries and time, conditioning on the country-specific crash risk p_i^I:

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

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

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

Conditional upon the country-specific crash risk p_i^I, the global crash risk parameters do not enter the likelihood:

$$f(Y_{i,\tau}|p_i^I, \alpha, \beta) = f(Y_{i,\tau}|p_i^I) \propto (p_i^I)^{Y_{i,\tau}\mathbf{1}^T}(1-p_i^I)^{\tau - Y_{i,\tau}\mathbf{1}^T}.$$

The joint posterior distribution of the parameters of interest then follows from Bayes' rule:

$$
\begin{aligned}
f(\{p_i^I\}_{i=1}^n, \alpha, \beta | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) &\propto f(\{D_{j,t}\}_{j=1,t=1}^{n,\tau} | \{p_i^I\}_{i=1}^n, \alpha, \beta) f(\{p_i^I\}_{i=1}^n | \alpha, \beta) f(\alpha, \beta) \\
&= \prod_{i=1}^n (p_i^I)^{Y_{i,\tau}\mathbf{1}^T} (1 - p_i^I)^{\tau - Y_{i,\tau}\mathbf{1}^T} \prod_{i=1}^n (p_i^I)^{\alpha-1} (1 - p_i^I)^{\beta-1} (\alpha + \beta)^{-5/2} \\
&= \prod_{i=1}^n (p_i^I)^{Y_{i,\tau}\mathbf{1}^T + \alpha - 1} (1 - p_i^I)^{\tau - Y_{i,\tau}\mathbf{1}^T + \beta - 1} (\alpha + \beta)^{-5/2}. \qquad (5)
\end{aligned}
$$

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 the country-specific information, that is, $f(p_i^I | \alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = f(p_i^I | \alpha, \beta, Y_{i,\tau})$.

$$
p_i^I | \alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau} \sim \text{Beta}(Y_{i,\tau}\mathbf{1}^T + \alpha, \tau - Y_{i,\tau}\mathbf{1}^T + \beta) \quad \forall i \qquad (6)
$$

One may think of $\alpha + \beta$ as the size of a pseudo sample the investor has at the back of her mind before observing the real data. α represents the number of market crashes in the sample while β represents the absence of crashes. Upon receiving new information, the investor simply add new occurrences of market crashes $Y_{i,\tau}\mathbf{1}^T$ to her existing observations α, and similarly for no-crash. Therefore, the investor updates her belief via an intuitive counting exercise.

However, in the model, the investor does not know the parameters α and β, as they depend on country-specific crash risk:

$$
f(\alpha, \beta | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = \frac{f(\{p_i^I\}_{i=1}^n, \alpha, \beta | \{D_{j,t}\}_{j=1,t=1}^{n,\tau})}{f(\{p_i^I\}_{i=1}^n | \alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau})}. \qquad (7)
$$

The investor thus needs to integrate out the uncertainty over the global crash risk parameters, to arrive at the country-specfic crash risk estimate:

$$
f(p_i^I | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = \int_{\alpha,\beta} f(p_i^I | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \alpha, \beta) f(\alpha, \beta | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) d(\alpha, \beta) \quad \forall i \qquad (8)
$$

Equation 7 and 8 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 7. We can interpret the mean of the common prior, $\alpha/(\alpha+\beta)$, as the global mean crash risk, that is the crash probability of an average country. Presumably, the investor would learn that the global crash risk follows a tight distribution, as the frequency of market crashes varies between 0% and 36% among different countries. The effective sample size, $\alpha+\beta$, controls the spread of the Beta distribution, with the dispersion in global crash risk decreasing in $\alpha+\beta$. Given that the investor can learn about the global crash risk distribution from the pooled cross-section, the posteriors for α and β have high precision.

The global crash risk parameters (α and β) further feed into the estimation of the country-specific crash risk in Equation 8. After we integrate out uncertainty with respect to these global crash risk parameters, we arrive at the marginal distribution for country-specific crash probability. The first component of the integrand is the conditional posterior of the crash probability for country i shown in Equation 6, which peaks around $(Y_{i,\tau}\mathbf{1}^T+\alpha)/(\alpha+\beta+\tau)$. This is the posterior mean one would arrive at by Bayesian updating according to the common prior and the observation that $Y_{i,\tau}\mathbf{1}^T$ crashes occurred in τ periods. Alternatively, one may write the posterior mean of country-specific crash risk p_i^I as

$$\mathbb{E}(p_i^I|\alpha,\beta,\{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = \frac{\frac{Y_{i,\tau}\mathbf{1}^T}{\alpha+\beta}+\frac{\alpha}{\alpha+\beta}}{1+\frac{\tau}{\alpha+\beta}}. \tag{9}$$

The subjective crash risk p_i^I increases in the number of crashes in country i observed up till time τ, $Y_{i,\tau}\mathbf{1}^T$. 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^I is shrunk towards the global mean crash risk, $\alpha/(\alpha+\beta)$.

$$\frac{\partial\,\mathbb{E}(p_i^I|\alpha,\beta,\{D_{j,t}\}_{j=1,t=1}^{n,\tau})}{\partial(\alpha+\beta)} = \frac{-Y_{i,\tau}\mathbf{1}^T + \frac{\alpha}{\alpha+\beta}\tau}{(\alpha+\beta)^2(1+\frac{\tau}{\alpha+\beta})^2} \begin{cases} \geq 0 & \text{if } \frac{Y_{i,\tau}\mathbf{1}^T}{\tau} \leq \frac{\alpha}{\alpha+\beta} \\ < 0 & \text{if } \frac{Y_{i,\tau}\mathbf{1}^T}{\tau} > \frac{\alpha}{\alpha+\beta} \end{cases} \tag{10}$$

Thus, when the global crash risk is estimated precisely, riskier countries get their posterior means scaled down while less risky countries get their posterior means scaled up. This is essentially the shrinkage effect achieved in the hierarchical Bayesian model.

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 its asymptotic behavior.

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

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^I integrates out uncertainty with respect to the global crash risk distribution, as in Equation 8. The resulting mean country-specific crash risk becomes a weighted average of the conditional means if global crash risk is observed. That is,

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

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 have assumed conditional independence of market crashes across countries. In the following sections we extend the model to account for conditional dependence.

C.2. Correlated Market Crashes

Recall the setup of the state-space problem from Equations 1 through 4. The presence of the systematic risk generates concurrence of market crashes across countries. To illustrate how the investor infers the magnitude of the systematic risk, we assume that its magnitude is constant over time and that the investor observes the latent states of nature.

Now that the idiosyncratic and systematic crash risk jointly determine the occurrence of market crashes, the data likelihood follows

$$f(\{D_{j,t}\}_{j=1,t=1}^{n,\tau} | \{p_i^I\}_{j=1}^{n}, p^S, \alpha, \beta, p_{HL}, p_{LH}) \propto \prod_{j=1}^{n} \prod_{t=1}^{\tau} (p_j^I + p^S \mathbf{1}_{x_t=H})^{D_{j,t}} (1 - p_j^I - p^S \mathbf{1}_{x_t=H})^{1-D_{j,t}}.$$

The detailed derivations of the marginal posteriors of the parameters of interest are given in Appendix B. We highlight the intuition behind the parameter identification. The investor uses cross-sectional information – how many markets crash simultaneously – to infer the magnitude of the systematic risk. On the other hand, idiosyncratic market crashes in the low risk regimes inform the degree of idiosyncratic risk in each market. A similar shrinkage effect, discussed in Section II.C.1, disciplines the inferred idiosyncratic risk. Overall, the model exploits maximum information in both dimensions to correct for survivorship bias in individual markets and account for an uncertain amount of systematic risk.

C.3. Bayesian Filter

The previous section assumes that the investor knows when the systematic regime starts and ends. The assumption simplifies the learning problem. In real life however, the difference between systematic and non-systematic states may not be so clear-cut. For more flexibility, we pursue a full Bayesian approach that incorporates this uncertainty. To better fit the data and the empirical observation that global financial markets are becoming increasingly integrated, we also allow the investor to infer the possibility of a structural break in the magnitude of the systematic risk (see Equation 3).

The likelihood conditional on the current information set I_t, the parameters of interest $\boldsymbol{\theta}$, and the latent state x_{t+1} takes the form

$$
\begin{aligned}
& p(\{D_{j,t+1}\}_{j=1}^{n}|I_t, \boldsymbol{\theta}, x_{t+1}) \\
\propto & \prod_{j} (p_j^I + (p_1^S \mathbf{1}_{t+1<t_b} + p_2^S \mathbf{1}_{t+1\geq t_b})\mathbf{1}_{x_{t+1}=H})^{D_{j,t+1}}(1 - p_j^I - (p_1^S \mathbf{1}_{t+1<t_b} + p_2^S \mathbf{1}_{t+1\geq t_b})\mathbf{1}_{x_{t+1}=H})^{1-D_{j,t+1}}.
\end{aligned}
$$

After the uncertainties over the latent states are integrated out, the complete data likelihood follows from the Markov property:

$$
p(\{D_{j,t}\}_{j=1,t=1}^{n,\tau}|\boldsymbol{\theta}) = \prod_{t=0}^{\tau-1} p(\{D_{j,t+1}\}_{j=1}^{n}|I_t, \boldsymbol{\theta}) = \prod_{t=0}^{\tau-1} \sum_{x_{t+1}} p(\{D_{j,t+1}\}_{j=1}^{n}|I_t, \boldsymbol{\theta}, x_{t+1}) p(x_{t+1}|I_t, \boldsymbol{\theta}).
$$

So far we have detailed the procedures of learning about the magnitude of the systematic risk and the correction for survivorship bias in the idiosyncratic crash risk. The following paragraphs discuss how the investor infers the states of the world and the location of the structural break.

Inference about the states of the world boils down to a filtering problem, which can be carried out in two steps – prediction and updating.

Prediction. Given the current period's posterior belief of being in the high risk p_t^H, the investor predicts next period's state of the world by accounting for the transition kernel and the uncertainty with respect to the current state.

$$
\begin{aligned}
p(x_{t+1} = H | I_t, \boldsymbol{\theta}) &= \sum_{x_t} p(x_{t+1} = H | I_t, \boldsymbol{\theta}, x_t) p(x_t | I_t, \boldsymbol{\theta}) \\
&= p_{LH} p_t^L + p_{HH} p_t^H \\
p(x_{t+1} = L | I_t, \boldsymbol{\theta}) &= \sum_{x_t} p(x_{t+1} = L | I_t, \boldsymbol{\theta}, x_t) p(x_t | I_t, \boldsymbol{\theta}) \\
&= p_{LL} p_t^L + p_{HL} p_t^H
\end{aligned}
$$

Updating. Upon observing market crashes in the next period, the investor updates on the probability of being in the high/low risk state.

$$
\begin{aligned}
p_{t+1}^H = p(x_{t+1} = H | I_{t+1}, \boldsymbol{\theta}) &= \frac{p(\{D_{j,t+1}\}_{j=1}^n, x_{t+1} = H | I_t, \boldsymbol{\theta})}{p(\{D_{j,t+1}\}_{j=1}^n | I_t, \boldsymbol{\theta})} \\
&= \frac{p(\{D_{j,t+1}\}_{j=1}^n | x_{t+1} = H, I_t, \boldsymbol{\theta}) p(x_{t+1} = H | I_t, \boldsymbol{\theta})}{\sum_{x_{t+1}} p(\{D_{j,t+1}\}_{j=1}^n | x_{t+1}, I_t, \boldsymbol{\theta}) p(x_{t+1} | I_t, \boldsymbol{\theta})} \qquad (11) \\
p_{t+1}^L = p(x_{t+1} = L | I_{t+1}, \boldsymbol{\theta}) &= \frac{p(\{D_{j,t+1}\}_{j=1}^n, x_{t+1} = L | I_t, \boldsymbol{\theta})}{p(\{D_{j,t+1}\}_{j=1}^n | I_t, \boldsymbol{\theta})} \\
&= \frac{p(\{D_{j,t+1}\}_{j=1}^n | x_{t+1} = L, I_t, \boldsymbol{\theta}) p(x_{t+1} = L | I_t, \boldsymbol{\theta})}{\sum_{x_{t+1}} p(\{D_{j,t+1}\}_{j=1}^n | x_{t+1}, I_t, \boldsymbol{\theta}) p(x_{t+1} | I_t, \boldsymbol{\theta})} \qquad (12)
\end{aligned}
$$

Note that p_{t+1}^H & p_t^H, and p_{t+1}^L & p_t^L each follow a recursion, intermediated by the regime switch probabilities and the conditional data likelihood.

To better understand how the investor learns about the location of the structural break t_b, we denote $\boldsymbol{\theta}^-$ as all the parameters in the model other than t_b.

Given an independent prior $p(t_b)$ over the location of the break point t_b, the joint likelihood of the data and the break point follows

$$
p(\{D_{j,t}\}_{j=1,t=1}^{n,\tau}, t_b | \boldsymbol{\theta}^-) = p(\{D_{j,t}\}_{j=1,t=1}^{n,\tau} | \boldsymbol{\theta}^-, t_b) p(t_b | \boldsymbol{\theta}^-) = p(\{D_{j,t}\}_{j=1,t=1}^{n,\tau} | \boldsymbol{\theta}^-, t_b) p(t_b).
$$

Summing over all possible locations of the structural break, we arrive at the likelihood of the data conditional on all other parameters. We can infer $\boldsymbol{\theta}^-$ in a similar manner as done in the previous section. The only difference is that we are integrating out uncertainty over the regime and uncertainty over the location of the structural break:

$$p(\{D_{j,t}\}_{j=1,t=1}^{n,\tau}|\boldsymbol{\theta}^{-}) = \sum_{t_b=1}^{\tau} p(\{D_{j,t}\}_{j=1,t=1}^{n,\tau}, t_b|\boldsymbol{\theta}^{-}). \tag{13}$$

Therefore, the conditional posterior of the structural break taking place at time t follows

$$p(t_b = t|\boldsymbol{\theta}^{-}, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = \frac{p(\{D_{j,t}\}_{j=1,t=1}^{n,\tau}, t_b|\boldsymbol{\theta}^{-})}{p(\{D_{j,t}\}_{j=1,t=1}^{n,\tau}|\boldsymbol{\theta}^{-})} \propto p(\{D_{j,t}\}_{j=1,t=1}^{n,\tau}, t_b|\boldsymbol{\theta}^{-}). \tag{14}$$

The marginal posterior of t_b can be obtained by integrating out uncertainty with respect to $\boldsymbol{\theta}^{-}$, where the posterior of $\boldsymbol{\theta}^{-}$ given the data can be derived from the data likelihood in Equation 13 and the prior:

$$p(t_b = t|\{D_{j,t}\}_{j=1,t=1}^{n,\tau}) = \int_{\boldsymbol{\theta}^{-}} p(t_b = t|\boldsymbol{\theta}^{-}, \{D_{j,t}\}_{j=1,t=1}^{n,\tau})p(\boldsymbol{\theta}^{-}|\{D_{j,t}\}_{j=1,t=1}^{n,\tau})d\boldsymbol{\theta}^{-}.$$

By drawing samples from the joint posterior distribution of all other parameters and propagating them through $p(t_b = t|\boldsymbol{\theta}^{-}, \{D_{j,t}\}_{j=1,t=1}^{n,\tau})$, one can simulate the posterior distribution of the break point.

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 survivor-bias-free sample starts in 1920 and ends in March 2020, thereby spanning a century. The choice of starting time is undoubtedly influenced by data availability. But more importantly, it is in the 1920s that there is large influx of American investors into the stock markets. It is suggested that this sudden boom in market participation is due to the near 20% inflation

American households experienced during World War I. Realizing that the US dollar no longer retains its value in bank deposits, investors showed growing interests in equity investments. We therefore think 1920 is a good starting point to model a US investor who learns about crash risk in the global equity markets.

Our sample covers a total of 55 countries. This includes all the MSCI developed markets, almost all the emerging markets, and several selected 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 (including dividends). As 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 US investor would earn by investing in global equity markets. We provide the summary statistics in Table I.

Market segmentation is substantial for most of the 20th century, which implies that a US investor cannot freely invests 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 overwhelming failure of absolute purchasing power parity, the notion of relative purchasing power parity, which states that countries with high inflation should see its currency depreciating, is accepted more broadly. Since we focus on the belief formation of a representative US investor, rather than the actual implementation of such trading strategies, we believe the usage of a common currency denomination is innocuous 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 US 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 that we do not have data for. Examples are Argentina prior to 1946 and Brazil prior to 1953. In such cases, we argue that data are more likely to be missing in bad times than good times. Hence, the missing data contain potentially a higher crash frequency than what we observe in-sample. Since the US crash risk is our primary focus and there is little doubt about the reliability of the US data, the selection issue in other markets can only bias downward our survivorship-corrected US 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, not every country is as lucky as the US in having 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 lived under the Communist regime, 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 the 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 have/had lived under the communist regime. Bialkowski and Ronn (2015) argue that the residual value of equity following major socio-economic disruptions are overstated in Goetzmann and Jorion (1999). Instead, they argue that severe restriction 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[7] in assuming a -100% return for stock markets in years when such 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 the investors.

Given our goal of studying crash risk, we identify a market crash as an annual return below -30%. It is the frequency of such event the investor is learning about and correcting survivorship bias for. We choose not to model learning about the magnitude of the crash 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 return expectation by aggregating information from the cross-section.

Finally, for the model in Section II.C.2 which augments the investor's information set with the nature of the state she is in, we identify high systematic risk states as those in which more than 20% of countries in the cross-section experience a market crash. Such periods include 1931 (the Great Depression), 1947 (the aftermath of World War II), 1974 (the collapse of Bretten Woods and the oil crisis), 1981 (global recession), 2008 (Global Financial Crisis), and 2020 (COVID-19). Both the intensive (the duration of risky periods) and the extensive margin (the likelihood of moving into a risky period) pin down the regime switch probabilities.

[7]See Table 7.

B. *Adjustment for unequal sample lengths*

Since the return series differ in length, we adjust for unequal sample length in the spirit of Stambaugh (1997).[8] In particular, we adjust for the first and second order moments of returns by regressing time series of shorter histories on those with full histories. The intuition is as follows. Suppose that two return series are positively correlated in the periods they overlap, and that the mean of the longer series is higher than its sample mean from the overlapping periods, then the unconditional average return of the shorter series should also be higher than its observed sample mean. In other words, the correlation structure informs the values of missing data. Similarly, we adjust for the variance of the shorter series. The tables and figures presented in the paper use the adjusted first and second order moments throughout.

C. *Method*

Due to the departure of the posterior from standard statistical distributions, estimating the model is non-trivial. For the most complicated filtering problem, the investor is interested in a total of 63 parameters: p_i^I for each of the 55 markets' idiosyncratic crash risk, the systematic risk pre and post break p_1^S and p_2^S, the parameters governing the idiosyncratic crash risk distribution α and β, the regime switch probabilities p_{LH} and p_{HL}, and the location of the structural break t_b. However, the good news is that we know the target density is likely to be uni-modal, as the posterior crash risk tends to concentrate around a small value. This alleviates concerns over the ergodicity of the MCMC algorithm. To simulate the posterior belief of the investor, we adopt one of the most recent advancements in MCMC – Hamiltonian Monte Carlo (HMC).

The algorithm improves upon those commonly used in the literature with a big gain in efficiency. A standard Metropolis-within-Gibbs sampler can theoretically work in our framework. However, the Gibbs sampler relies on sampling iteratively from the conditional

[8] See Proposition 1.

distribution. In a simple bivariate environment, this implies that the sampler can only move in the horizontal or the vertical direction within one iteration, without efficiently moving along the diagonal if the two parameters are correlated. The high dimentionality of the problem at hand makes it ill-suited for our purpose. By contrast, HMC exploits gradient information of the target density to optimally choose the direction to move in. To do so, it introduces an auxiliary variable called momentum. The algorithm generates a new sample from a probability distribution by giving the point some momentum and then updating the position and momentum according to a certain system of differential equations (Hamilton's equations, hence the name). The sampler then follows the Metropolis-Hastings algorithm in accepting or rejecting the new proposed value. For illustration, one may think of the target density as a bowl with a frictionless surface. We flick a particle on this surface in a random direction and let it travel across the surface for some time before flicking in a new random direction. By repeating the process, the particle traces out the shape of the surface and hence the posterior distribution. We refer interested readers to Neal (2011) for a rigorous explanation of the algorithm.

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[9], where the first 2,000 draws serve as burn-ins[10]. 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[11] to ensure the convergence of the algorithm.

[9] 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.

[10] This is a slight abuse of language because they are technically not burn-ins, but a warm-up period for tuning parameter adaption.

[11] 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.

IV. Model Results

A. Idiosyncratic Crash Risk

We estimate the model in Section II.C.1, which assumes that crash realizations are conditionally independent across time and markets. The posterior summary statistics can be found in the third column in Table III.

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) based on the full sample. 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 6.43% (posterior mean). That is, the investor would correct the survivorship bias in terms of crash risk in the US sample by 28.6%. The strongest shrinkage effect is that for Turkey: the posterior mean crash risk is shrunk by almost a 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 on, 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-sectional to discipline the perceived crash risk for Morocco because of a lack of sufficient observations. As a results, 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. The interesting phenomenon is exactly due to the unconventional 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 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 9. 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 10.

What is an average market in the sample? Figure 4 reveals that it is Hong Kong, of which the subjective crash risk and the in-sample frequency are both around 9.5%. 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, it is certainly not the one in the 1950s, since the stock market did not exist back then. Ultimately, it is the data that decide which market has the average crash risk in the cross-section.

US v.s. An Average Country. With an uninformative prior over the global crash risk distribution, the investor forms crash belief with an expanding sample. Figure 5 plots the subjective crash risk for the US and the average country (global mean crash risk) from 1921 and 2020. We omit the common prior mean of 50% for ease of reading. 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. Typically, investor's subjective crash risk is right-skewed[12]. When faced with uncertainty about crash risk, the investor puts more weight on high crash risk.

First, we observe that the global mean crash risk quickly stabilizes around 9.25%. The abundance of data pooled from the cross-section allows a more precise estimate of the global

[12]This could be an artifact of our assumption of a common prior, since the Beta distribution has skewness built in. Nevertheless, neither the unconditional posterior of the country-specific crash risk nor the global mean crash risk follow a Beta distribution exactly.

mean crash risk compared to the country-specific ones. The subjective crash risk in the US 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 US 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 US departs steadily away from that the global mean, due to the absence of negative shocks in the US sample. We observe that the spikes in subjective crash risk in 1974, 2008 and 2020 move in synchronicity with those in the global mean crash risk, as they are all systematic shocks. Combined with the fact the posterior estimate is becoming more precise, the adjustments in subjective crash belief are attenuated. As of March 2020, there is a 2.82% difference between the crash risk in the US (6.43%) and the global average (9.25%).

B. Systematic Crash risk

Clearly, systematic and idiosyncratic risk have different effects on the investor's subjective crash belief. We now estimate the model in Section II.C.2 and provide the posterior estimates in the fourth column of Table III. The subjective idiosyncratic crash risk in the US is 2.12%. The systematic risk accounts for the difference between this estimate and the one in the previous section. In a risky period, the likelihood for any country to experience a crash is increased by 39.26%. Moreover, the world has a 9.76% chance of switching into a risky regime. Once in a risky regime, normal times take over with probability 90.87%. It implies that the world is in a normal state 90.3% of the time and in a risky state 9.7% of the time. The unconditional probability of a market crash in the US is therefore 5.9%. This is slightly lower than the posterior crash belief she would arrive at by assuming that crash risk is uncorrelated across countries.

Cross-Sectional Comparison. We plot in Figure 6 the idiosyncratic crash risk (in blue), p_i^I, and the unconditional crash risk (in red), $p_H(p^S+p_i^I)+p_L p_i^I$, for each country against the historical frequency. The inference for idiosyncratic crash risk is influenced by how often the market crashes outside of the high risk regimes. For most countries, the idiosyncratic crash risk estimates are lower than those in the previous section because the common variation in crashes taken place in risky periods is absorbed by the systematic crash risk parameter. On the other hand, the unconditional crash risk in each country is higher than the idiosyncratic crash risk by the same proportion, because countries have the same exposure to the systematic risk.

Evolution Of Systematic Risk. To get a better idea of the extent of the systematic risk, we plots its evolution in the bottom panel in Figure 7. At the start of sample, the posterior mean stays at the uninformative prior mean of 50% with a wide credible set because the world has not experienced any systematic shock and therefore there is no learning about the systematic risk. Upon the arrival of the Great Depression in 1931, the posterior mean systematic risk dips to 45% and the credible set quickly shrinks. This is in spite of the fact that the Great Depression is one of the most systematic events in the sample, affecting more than half of the countries in the sample. The reason behind the finding is that the idiosyncratic crash risk estimates are still high at the start of the learning, which crowds out the estimation of systematic risk. The high systematic risk periods following World War II affected a limited number of countries, resulting in a sharp decline in the systematic crash risk. It is not until when the Global Financial Crisis hits in 2008 that the systematic risk is boosted up to 39.3%.

US v.s. An Average Country. The top two panels in Figure 7 shows the evolution of crash belief for the US and the average country. Given the systematic nature of the market crashes in 1974, 2008 and 2020, the investor no longer adjusts upwards her belief about the US crash risk in these periods. Compared with Figure 5, the divergence of the US crash risk (2.12%) from the global mean (5.47%) is much more significant.

To summarize, the model that considers systematic crash risk manages to correct for survivorship bias in the US sample by 34.8%. The subjective crash belief for the US is distinct from that of an average country.

C. Bayesian Filter

We simulate the belief formation of an investor according to the filtering algorithm discussed in Section II.C.3. The filter is carried out with an expanding information set through time. Prior to 2008, the investors fails to infer a structural break in the data. However, once the sample extends up to and beyond 2008, the location of the structural is accurately learnt. Implicitly, a long-lived investor also faces model uncertainty in inferring crash risk. We allow her the freedom to fitting two models to the data and choosing the better one to form posterior belief.[13]

Table III summarizes the posterior estimates for the parameters of interest in the last column. The global mean idiosyncratic crash risk is 4.9% while that for the US is 1.8%. The probability of switching into the high risk regime is estimated to be 12.6% while the probability of returning to a low risk regime is 66.7%. Overall, it implies the world would be in a high systematic risk regime 16% of the time. This is roughly in line with our experience in the past 100 years, where 14 years are identified to be high systematic risk states. More interestingly, the magnitude of systematic risk is estimated to be 24.7% (p_S^1) prior to the structural break and 80.4% (p_S^2) post the structural break. Both parameters are precisely estimated, evidenced by the tight posterior credible sets. Recall that we have independent uniform priors over the magnitude of the systematic risk pre and post structural break. Therefore, the divergence of the posterior from the prior strongly suggests a structural change in the data.

Where Is The Structural Break? The bottom panel in Figure 8 plots the probability that a structural point takes place at each point in time (see Equation 14) using the full sample. The probabilities sum to 1 and are close to 0 most of the time. It is only from 2001 to 2008 that the posterior probability is significantly positive. In fact, the peak is reached in 2008, hence identifying the point of the structural break.

[13]Our motivation to include a structural break in the analysis arises from the failure of a simple filter without structural breaks to fit the data, as the HMC algorithm fails to converge once the sample extends beyond 2008.

States Of The World. The top panel in Figure 8 shows the probability that the world is in a high risk state, where crash risk increases by p_t^S in every country. Years in which such probability is above 0.5 include 1931, 1940, 1947-1949, 1974, 1981-1983, 1998, 2000, 2008, 2011, and 2020, covering periods with significant global shocks. In addition, the filter produces a moderate level of persistence in the high systematic risk states. This differs from the hard threshold rule used in the previous section in classifying systematic states of the world, resulting in little persistence in high systematic risk states. The persistence in risky states is particularly detrimental to investors and has important asset pricing implications.[14]

Evolution Of Systematic Risk. The magnitude of the systematic risk varies over time (Figure 9). It spikes to 50% during the Great Depression but subsequently stabilizes around 25% during most of the 20th century. However when 2008 comes around, the magnitude of the systematic risk reaches a high plateau of around 80% and falls off slightly in 2011 and 2020. Due to the limited length of sample since the occurrence of the structural break, the question remains that is the increase in systematic risk permanent or transitory? The non-stationary feature of the data might also represent a third regime with an ultra-high level of systematic risk in a stationary world. Nevertheless, a three-stage regime switch model fails to fit the data,[15] likely because the the investor fails the infer the probability of exiting the ultra-high systematic risk regime.

Cross-Sectional Comparison. Figure 10 plots the idiosyncratic crash risk p_i^I, and the unconditional crash risk $p_H(p_2^S + p_i^I) + p_L p_i^I$, against the historical frequency of market crashes based on the full sample. The unconditional crash risk is much higher than the idiosyncratic crash risk because a structural break took place in the magnitude of systematic risk. Together with the modest estimate that the world is in a high risk regime 16% of the time, the systematic risk dominates and the investor infers a high unconditional crash risk for all countries. In contrast, we also plot in Figure 11 the posterior means at 2007, before the structural break took place. Given the shorter sample, the shrinkage effect on idiosyncratic

[14] The point is emphasized by Collin-Dufresne et al. (2016).

[15] The HMC sampling scheme did not converge.

crash risk is stronger. Meanwhile, the pre-structural break value of the systematic risk yields lower unconditional crash risk estimates for all countries.

US v.s. An Average Country. Figure 12 highlights the secular decline of the subjective crash risk in the US from that of an average country. The growing divergence is the most apparent because the filter picks up all the systematic risk and assign it to the high-risk regime. Although the posterior crash belief for every country should follow a martingale, the non-stationarity is accentuated for the US. We discuss in the following section the asset pricing implication for the secular decline.

V. Asset Pricing Implication

A. Declining Equity Premium

In order to draw asset pricing implication 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 potentially 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. Existing literature identifies the rare occurrence of such events from output and consumption growth disasters around the world. And the disaster risk models calibrate the consumption growth process in the US with international observations of disasters, which are of a higher frequency. However as we show in the introduction, the international equity premium estimated from a global CAPM coincides with that estimated based on the US experience. In that sense, the disaster risk literature is implicitly using international data to calibrate an international equity premium.

To draw a parallel between crash risk and disaster risk, we make use of the finding in Nakamura et al. (2013) that 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 belief on crash risk to belief on 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 under imperfect information with her current posterior estimate of disaster risk, which maps to crash risk.

Consumption follows a random walk with drift and contains a disaster component. The log consumption has mean μ and standard deviation σ in periods without disasters.

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

A disaster v_{t+1} takes place with probability p, which is unknown to the agent. The agent has a subjective belief about disaster risk p_t, which gets updated each period. Consumption drops by b upon impact.

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

The representative consumer maximizes a time-additive isoelastic utility function, where γ is the relative risk aversion.

$$U_t = E_t[\sum_{s=t}^{\infty} \frac{C_s^{1-\gamma}}{1-\gamma}]$$

The risk premium on the consumption claim has the standard representation. The risk

[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.

premium rises in the subjective disaster risk p_t.

$$\log[\frac{E_t(1+R_{t+1})}{1+R_f}] = \gamma\sigma^2 + b[(1-b)^{-\gamma} - 1]p_t \tag{15}$$

Under the assumption that the disaster risk p_t follows a similar trend as crash risk, the subjective disaster risk p_t in the US should also experience a secular decline in the past century.[17] It implies that the equity premium in the US should decline accordingly, which is in part reflected in the ever-increasing valuation ratio. The decline in equity premium has been widely noted in the literature, e.g. Lettau et al. (2008). The pattern is furthermore corroborated by the survey data from the Survey of Professional Forecasters, where the forward-looking 10-year expected equity premium has experienced a secular decline since the 1990s from 5% to 3%. By contrast, the stability of global crash risk implies that there should be a growing divergence between the risk compensation in the global equity market and that in the US market. With a globally diversified portfolio, one may tame down the volatility in the international equity market while earning a sizable international equity premium. The Sharpe ratio on the global equity market may therefore be comparable, if not higher, than that for the US.

B. *Decomposing Of The Equity Premium*

Bayesian updating implies that any shock to the current posterior will be permanent, as the posterior means form a martingale. In a world with imperfect information, the agent that learns about disaster risk would experience a series of permanent shocks to her belief. In turn, the belief updating would result in a sequence of discount rate and cash flow shocks

[17]Is it sufficient to consider only the posterior mean crash (disaster) risk, but not the posterior uncertainty? We offer two justifications. The first is that the posterior mean is the optimal point estimate under a quadratic loss function. Secondly, the posterior mean would be sufficient if the agent has a Beta prior and observes crash realizations that follow a Bernoulli distribution. The predictive density $p(D_{t+1}=1) = \int p(D_{t+1} = 1|p)f(p|I_t)dp = \int p\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}dp = \frac{\alpha}{\alpha+\beta}$. This is not exactly true in our setup, because we allow the investor 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.

that manifest themselves in the historical returns. In this section, we quantify the upward bias in the US historical equity premium that is due to learning induced positive returns news and negative shocks that did not realize in-sample.

In an i.i.d. economy, the permanent shock to perceived disaster risk translates into permanent shocks to all future cash flows and discount rates. Suppose that the agent treats the current posterior mean as the true constant parameter and ignores any dynamic in the endogenous state transition that arise from learning, the valuation ratio is now time varying only from updating.[18] 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.

Consider the same setup of the problem as that in the previous section, where the representative agent has CRRA utility and the consumption growth process is i.i.d. with a disaster component. The price-consumption ratio, $Z_t^C = P_t^C/C_t$, in the economy depends on p_t, where

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

The expected return follows

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

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 a constant based on time t information. By contrast, the

[18] À la Weitzman (2007). This is otherwise known as anticipated utility (AU). First introduced by Cogley and Sargent (2008) as an approximation to the full Bayesian problem, AU is widely used when computation may be infeasible due to non-stationary evolution of the posterior.

ex-post realized return is

$$R_{t+1} = \frac{P^C_{t+1} + C_{t+1}}{P^C_t} = \frac{Z^C_{t+1} + 1}{Z^C_t} \frac{C_{t+1}}{C_t}$$

The realized return reflects changes in valuation ratio that are not taken into account from an ex-ante perspective. The difference between the expected and realized returns reflects the change in valuation ratio as well as cash flow surprise. To see it more clearly:

$$\mathbb{E}_t[R_{t+1}] = \frac{e^{\mu+\frac{1}{2}\sigma^2}[(1-b)p_t + 1 - p_t]}{K_t}$$
$$R_{t+1} = \frac{1-K_t}{1-K_{t+1}} \frac{1}{K_t} \frac{C_{t+1}}{C_t}$$

For a sample absent realizations of disasters, 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 excepted cash flow, $e^{\mu+\frac{1}{2}\sigma^2}[(1-b)p + 1 - p]$. In addition, if the valuation ratio rises following a decline in disaster risk, $\frac{1-K_t}{1-K_{t+1}} > 1$, the realized return would be higher than the expected return. However, in order for the valuation ratio to go in the right direction when disaster risk declines, we need that $\gamma < 1$. Otherwise, the rise in risk free rate more than offsets the decline in the risk premium, leading to a decline in the valuation ratio. If the agent is endowed with different willingness to smooth across time and across states, the rise in risk-free rate following a decline in disaster risk can be attenuated and the return surprises would be uniformly positive.

To illustrate this point, we price a dividend claim with Epstein-Zin-Weil preference. The dividend claim is a levered claim on consumption, such that $\Delta d = \lambda \Delta c$. The utility function takes the standard form, 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.

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

Let $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$. One can show that the pricing kernel in the economy is affine in the 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 equations.

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

The equilibrium valuation ratios are functions of the current subjective disaster risk.

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

One may verify that $Z_t^C = Z_t^D$ when the stock market is not levered and $\lambda = 1$, and that when $\gamma = \frac{1}{\psi}$, the price-consumption ratio is the same as that in the CRRA case. The expected and realized rates of return on the dividend claim equal

$$\mathbb{E}_t[R_{t+1}^D] = \mathbb{E}_t[\frac{Z_t^D + 1}{Z_t^D}\frac{D_{t+1}}{D_t}] = \frac{1 - K_t^D}{1 - K_t^D}\frac{1}{K_t^D}\mathbb{E}_t[\frac{D_{t+1}}{D_t}]$$
$$R_{t+1}^D = \frac{Z_{t+1}^D + 1}{Z_t^D}\frac{D_{t+1}}{D_t} = \frac{1 - K_t^D}{1 - K_{t+1}^D}\frac{1}{K_t^D}\frac{D_{t+1}}{D_t}$$

For the valuation ratio to be non-increasing in disaster risk, we need that $K_{t+1}^D > K_t^D$ 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 $\lambda > \gamma \geq \frac{1}{\psi}$.[19] 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 can see that the difference mainly arises from two channels: *luck* and *learning*.

$$R^D_{t+1} - \mathbb{E}_t[R^D_{t+1}] = \underbrace{\mathbb{E}_t[\frac{D_{t+1}}{D_t}](\frac{1-K^D_t}{1-K^D_{t+1}}\frac{1}{K^D_t} - \frac{1-K^D_t}{1-K^D_t}\frac{1}{K^D_t})}_{Learning} + \underbrace{\frac{1-K^D_t}{1-K^D_t}\frac{1}{K^D_t}(\frac{D_{t+1}}{D_t} - \mathbb{E}_t[\frac{D_{t+1}}{D_t}])}_{Luck}$$
$$+ (\frac{1-K^D_t}{1-K^D_{t+1}}\frac{1}{K^D_t} - \frac{1-K^D_t}{1-K^D_t}\frac{1}{K^D_t})(\frac{D_{t+1}}{D_t} - \mathbb{E}_t[\frac{D_{t+1}}{D_t}])$$

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

In addition, as the agent revises her perception of disaster risk downwards over the century, the ex-post price-dividend ratio similarly experiences positive revisions upwards. Although the agent ex-ante believes 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*. The residual term is negligible in magnitude. In Figure 13, we decompose the wedge between realized and expected rates of return[20] into the aforementioned *luck* and *learning* components. For the United States, the realized rate of return exceeds the expected rate of return by 2%, which is the amount the historical equity premium that has been overstated by. On a closer look, the 2% is about equally split

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

[20]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 expected and historical equity premium.

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 model early-on decides that Switzerland is a low-risk country. The decision attenuates the perception of luck and contributes to a steady downward revision in disaster risk. By contrast, the model ascribes most of Morocco's excess return to *luck* because it is deemed to be riskier than Switzerland.

To conclude, we note that the return news in the US has been predominately positive, despite that such news should be zero on average in theory. The secular decline in perceived disaster risk as opposed to a stable global disaster risk implies that the historical US equity premium is in part due to positive return surprises rather than compensation for risk. After all, our analysis suggests that the US equity premium puzzle has been overstated by one-third.

VI. Conclusion

Motivated by striking finding in Goetzmann and Jorion (1999) that the the US equity performance appears to be an exception rather than the norm, we investigate the validity of the equity premium puzzle in a comprehensive dataset on international equity returns. We find the 6% equity premium studied extensively in the literature is implicitly supported by the international evidence. The high volatility of equity returns in emerging markets implies a positive risk-return trade-off in the classical sense (using arithmetic return). Nevertheless, we note that global equity market does not yield a high buy-and-hold return on average. This suggests that investors in emerging markets are compensated for their ability to endure volatility on its portfolio and benefit from timing the market.[21] The fact that the United States and most advanced economies lie on or slightly below the security market line suggests

[21] For illustration, in Figure 3 we plot the buy-and-hold return of investing $1 in max{1920, when the market enters the sample} for selected countries in our sample. The smooth and continuous price history of the US is in sharp contrast to many other countries.

that the they are subject to little survivorship bias, as suggested by Li and Xu (2002) and Dimson et al. (2008).

We approach the question from a new perspective. We propose a learning model where the investor accounts for country-level heterogeneity in inferring the true crash risk in the US. The assumption that the crash risk in each country is drawn from the same distribution is plausible for an investor standing in 1920, since the global economies are relatively homogeneous at the point. With a hierarchical model, we account for the possibility that a US investor is being compensated for disastrous outcomes in the equity market which did not materialize in-sample. We show across three different models that the subjective crash risk in the US, after correcting for survivorship bias, is strictly higher than the historical frequency of such event. More importantly, the US experience gradually deviates from that of the global average. In addition, we show that a structural break in the magnitude of systematic risk took place at the turn of the century, with the model ascribing the highest possibility of a structural break to 2008. This accords with our experience of rising globalization.

We discuss the asset pricing implication of the findings in an equilibrium setting. First, The secular decline in crash risk implies a lower equity premium going forward. Second, the benefit of global diversification is reduced as global economies are becoming increasingly integrated. The downside co-movement looms large. Lastly, the US equity premium puzzle should in fact be a smaller puzzle, as the historical equity premium reflects that US 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. We suggest a 2% reduction of the target equity premium from 6% to 4% for future finance research.

Appendix A. Choice of Prior

In the section, we discuss the intuition behind the hyperprior. As explained before, we model an investor who starts learning about global crash risk in 1920. It is reasonable to assume that she has no knowledge over the parameters governing the global crash risk distribution. Recall that our common prior on idiosyncratic crash risk follows a Beta distribution: Beta(α, β). For clarity, consider the change of variables:

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

Now, we reinterpret the two degrees of freedom entailed in a Beta distribution. The mean of the Beta distribution ϕ is the average crash risk globally. The effective sample size λ controls the dispersion of the distribution. The distribution of crash risk is more dispersed with a low λ.

For the investor to have an uninformative hyperprior, it immediately follows that ϕ should be uniform between 0 and 1. As for λ, we model $\lambda^{-1/2}$ to be uniform between 0 and 1. Implicitly, we assume that the hyperprior's effective sample size is at least 1, but we do not restrict where the peak is. The unscaled λ, in turn, follows a Pareto distribution. Its density has a power tail and decays at a slower rate than an exponential distribution. An engaged reader might ask: why not set the prior on λ such that it is uniform on $[1, \infty]$. As it turns out, the posterior density would be improper, making it an invalid subjective belief.

Overall, the prior on the two parameters can be summarized as follows.

$$\phi \sim \text{Uniform}[0, 1]$$

$$\lambda \sim \text{Pareto}(1, 0.5)$$

Hence, the joint distribution of ϕ and λ is

$$f(\phi, \lambda) \propto \lambda^{-3/2}$$

The hyperprior on α and β is therefore

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

Appendix B. Posterior Derivation

We derive the marginal posterior for each parameter of interest from the model discussed in Section II.C.2. First, the regime switch probabilities can be learnt separately from other parameters.

$$f(p_{LH}|\{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau}, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) \sim Beta(1 + \sum_{t=2}^{\tau} \mathbf{1}_{\{x_{t-1}=L\&x_t=H\}}, 1 + \sum_{t=2}^{\tau} \mathbf{1}_{\{x_{t-1}=L\&x_t=L\}})$$

$$f(p_{HL}|\{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau}, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}) \sim Beta(1 + \sum_{t=2}^{\tau} \mathbf{1}_{\{x_{t-1}=H\&x_t=L\}}, 1 + \sum_{t=2}^{\tau} \mathbf{1}_{\{x_{t-1}=H\&x_t=H\}})$$

The joint posterior distribution of the rest of the parameters of interest now takes the form

$$\begin{aligned} &f(\{p_i^I\}_{i=1}^n, p^S, \alpha, \beta|\{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau}) \\ &\propto f(\{D_{j,t}\}_{j=1,t=1}^{n,\tau}|\{p_i^I\}_{i=1}^n, p^S, \alpha, \beta) f(\{p_i^I\}_{i=1}^n|\alpha, \beta) p(\alpha, \beta) f(p^S) \\ &= \prod_{i=1}^{n}\prod_{t=1}^{\tau}(p_i^I + p^S\mathbf{1}_{x_t=H})^{D_{i,t}}(1 - p_i^I - p^S\mathbf{1}_{x_t=H})^{1-D_{i,t}} \prod_{i=1}^{n}(p_i^I)^{\alpha-1}(1-p_i^I)^{\beta-1}(\alpha+\beta)^{-5/2} \end{aligned}$$

The systematic risk parameter is only informed by crash realizations in the high risk regimes. It follows a generalized version of a Beta distribution, where the fraction of countries experiencing market crashes in high systematic risk periods is key for identifying the extent of the systematic risk.

$$\begin{aligned} &f(p^S|\{p_i^I\}_{i=1}^n, \alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau}) \\ &\propto \prod_{i=1}^{n}\prod_{t|\mathbf{1}_{x_t=H}}(p_i^I + p^S)^{D_{i,t}}(1 - p_i^I - p^S)^{1-D_{i,t}} \\ &= \prod_{i=1}^{n}\left[(p_i^I + p^S)^{\sum_{t|\mathbf{1}_{x_t=H}} D_{i,t}}(1 - p_i^I - p^S)^{\sum_{t|\mathbf{1}_{x_t=H}}(1-D_{i,t})}\right] \end{aligned} \quad \text{(B1)}$$

The idiosyncratic crash risk is identified in a similar fashion as before. Common variation in high systematic risk periods is already accounted for by p^S. As a result, it is market crashes in normal times and the common prior that decide the posterior belief about idiosyncratic crash risk.

$$f(p_i^I|p^S, \alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})$$
$$\propto \prod_{t=1}^{\tau}(p_i^I + p^S\mathbf{1}_{x_t=H})^{D_{i,t}}(1 - p_i^I - p^S\mathbf{1}_{x_t=H})^{1-D_{i,t}}(p_i^I)^{\alpha-1}(1 - p_i^I)^{\beta-1} \tag{B2}$$

The joint conditional posterior of the idiosyncratic and systematic crash risk parameters can be written as

$$f(\{p_i^I\}_{i=1}^n, p^S|\alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})$$
$$\propto \prod_{i=1}^{n}\prod_{t=1}^{\tau}(p_i^I + p^S\mathbf{1}_{x_t=H})^{D_{i,t}}(1 - p_i^I - p^S\mathbf{1}_{x_t=H})^{1-D_{i,t}}\prod_{i=1}^{n}(p_i^I)^{\alpha-1}(1 - p_i^I)^{\beta-1}$$

Hence, we arrive at the following expression for the marginal posterior distribution for the global crash risk parameters.

$$f(\alpha, \beta|\{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau}) = \frac{f(\{p_i^I\}_{i=1}^n, p^S, \alpha, \beta|\{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})}{f(\{p_i^I\}_{i=1}^n, p^S|\alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})} \tag{B3}$$

We can integrate out uncertainty with respect to the global crash risk parameters in Equation B3 to derive the marginal posterior distribution of the idiosyncratic crash risk estimate p_i and the systematic crash risk p^S.

$$f(p^S, \{p_i^I\}_{i=1}^n|\{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})$$
$$= \int_{\alpha,\beta} f(p^S, \{p_i^I\}_{i=1}^n|\alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})f(\alpha, \beta|\{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})d(\alpha, \beta)$$
$$f(p_i^I|p^S, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})$$
$$= \int_{\alpha,\beta} f(p_i^I|p^S, \alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})f(\alpha, \beta|\{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})d(\alpha, \beta)$$
$$f(p^S|\{p_i^I\}_{i=1}^n, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})$$

$$= \int_{\alpha,\beta} f(p^S | \{p_i^I\}_{i=1}^n, \alpha, \beta, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau}) f(\alpha, \beta | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau}) d(\alpha, \beta)$$

$$f(p^S | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau}) = \frac{f(p^S, \{p_i^I\}_{i=1}^n | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})}{f(\{p_i^I\}_{i=1}^n | p^S, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})}$$

$$f(\{p_i^I\}_{i=1}^n | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau}) = \frac{f(p^S, \{p_i^I\}_{i=1}^n | \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})}{f(p^S | \{p_i^I\}_{i=1}^n, \{D_{j,t}\}_{j=1,t=1}^{n,\tau}, \{\mathbf{1}_{x_t=H}\}_{t=1}^{\tau})}$$

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.

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, 2015, Financial Markets in the Face of the Apocalypse, *SSRN Electronic Journal* .

Brown, Stephen J, William N Goetzmann, and Stephen A Ross, 1995, Survival 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* .

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).

Dimson, Elroy, Paul Marsh, and Mike Staunton, 2008, *The Worldwide Equity Premium: A Smaller Puzzle* (Elsevier B.V.).

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.

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.

Li, Haitao, and Yuewu Xu, 2002, Survival bias and the equity premium puzzle, *Journal of Finance* 57, 1981–1995.

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.

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

Stambaugh, Robert F, 1997, Analyzing Investments Whose Histories Differ in Length, *Journal of Financial Economics* 45, 285–331.

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.

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.

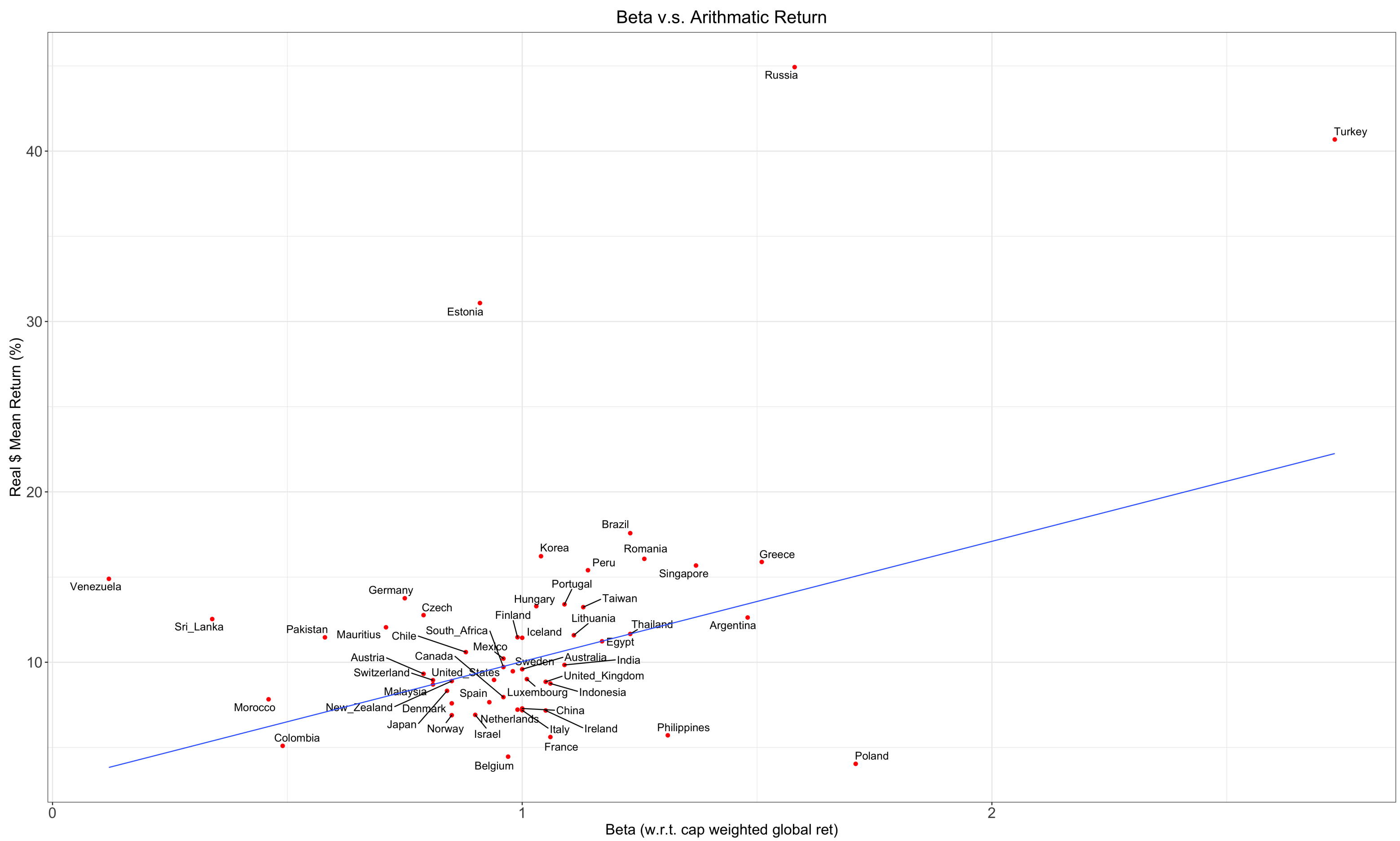


Figure 1. Global CAPM

The figure plots the real dollar mean return (adjusted for unequal sample lengths) against the exposure (beta) to a capitalization weighted global portfolio. The return series are at an annual frequency (year-end to year-end, except for 2020 in which the series end on March 23) and are converted to real US dollar to facilitate cross-country comparisons. The sample covers 1920-2020 and includes a total of 55 markets. The capitalization weighted global portfolio is constructed annually with existent markets in the cross-section. A median regression line (shown in blue) is fitted to the data, which has slope 7.1% (global equity premium) and intercept 3% (risk-free rate).

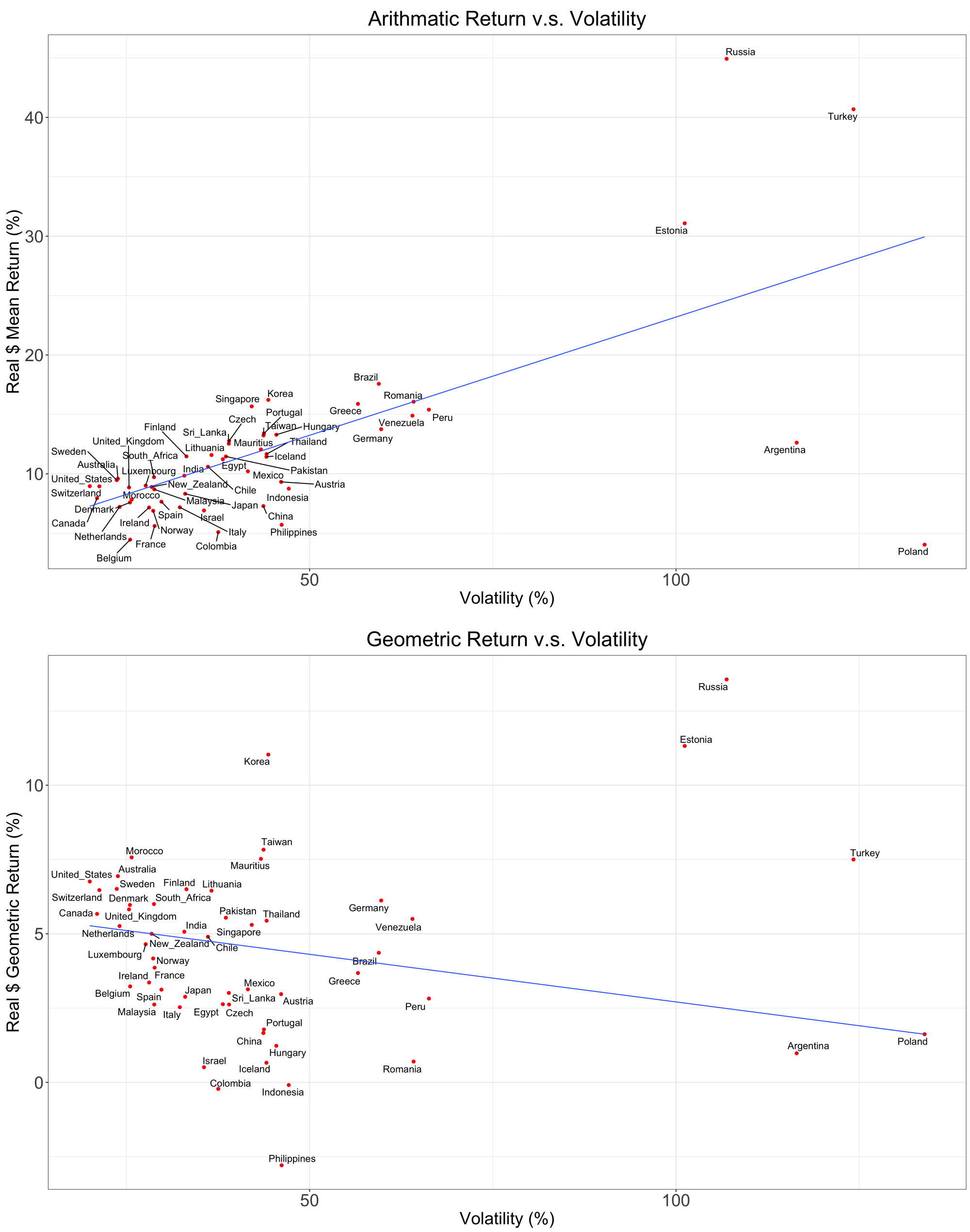


Figure 2. Risk Return Trade-off

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.

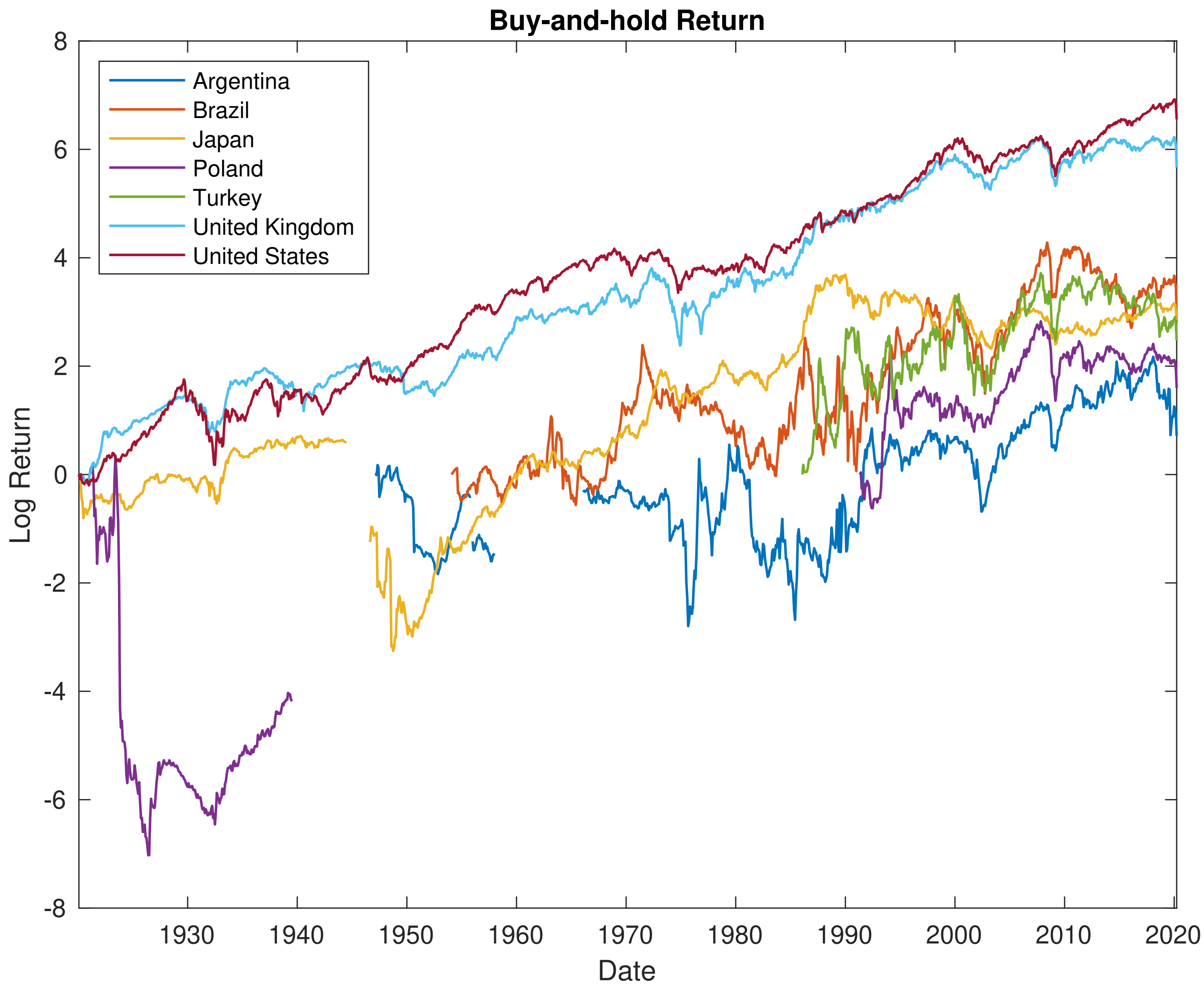


Figure 3. Buy-and-Hold Return

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 US stands out in having a smooth and continuous series, whereas other markets experience trading breaks and high volatility.

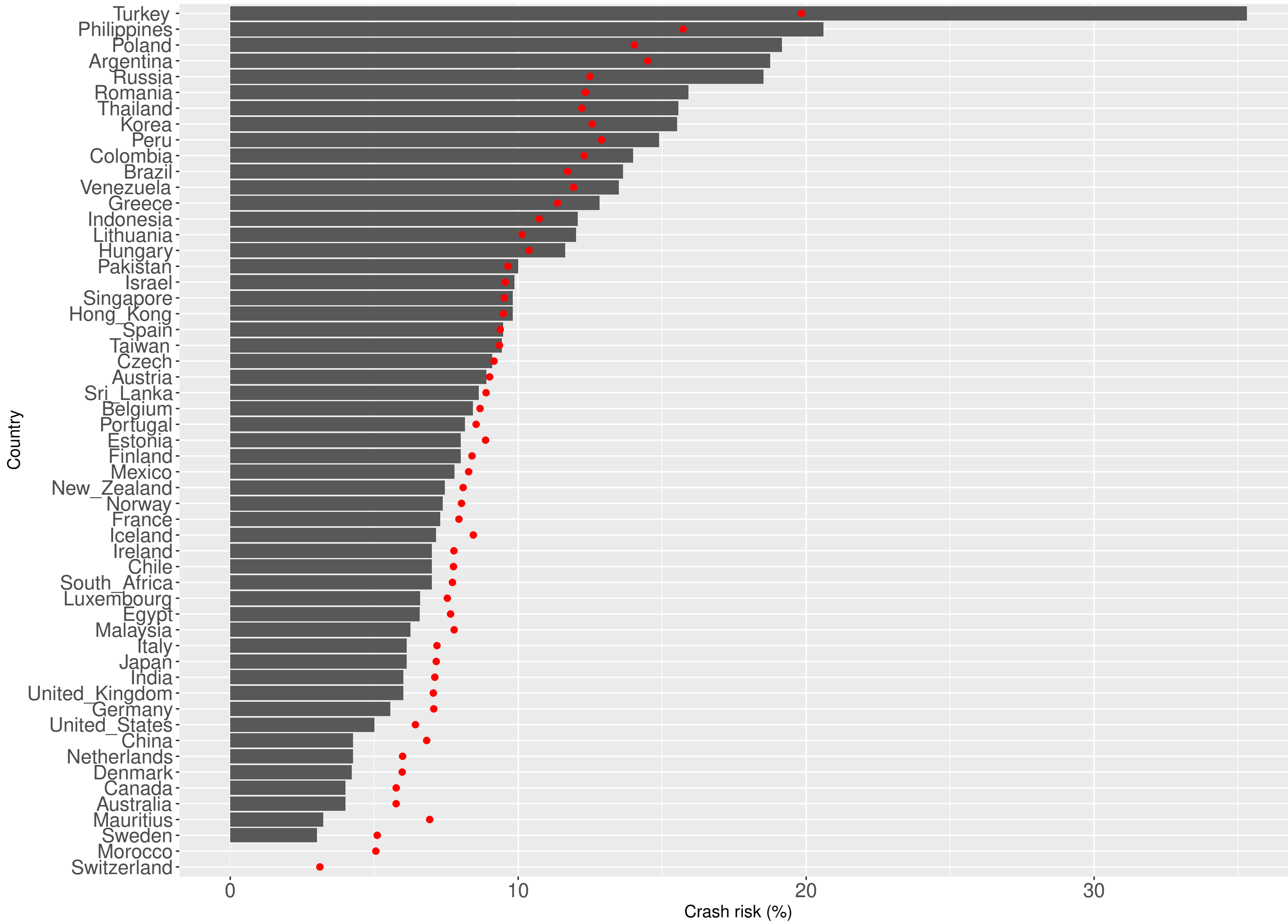


Figure 4. Subjective Crash Risk v.s. In-sample Frequency

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.C.1, with summary statistics in the third column in Table III. 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.C.1. 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 US crash risk is shown in the dashed blue line.

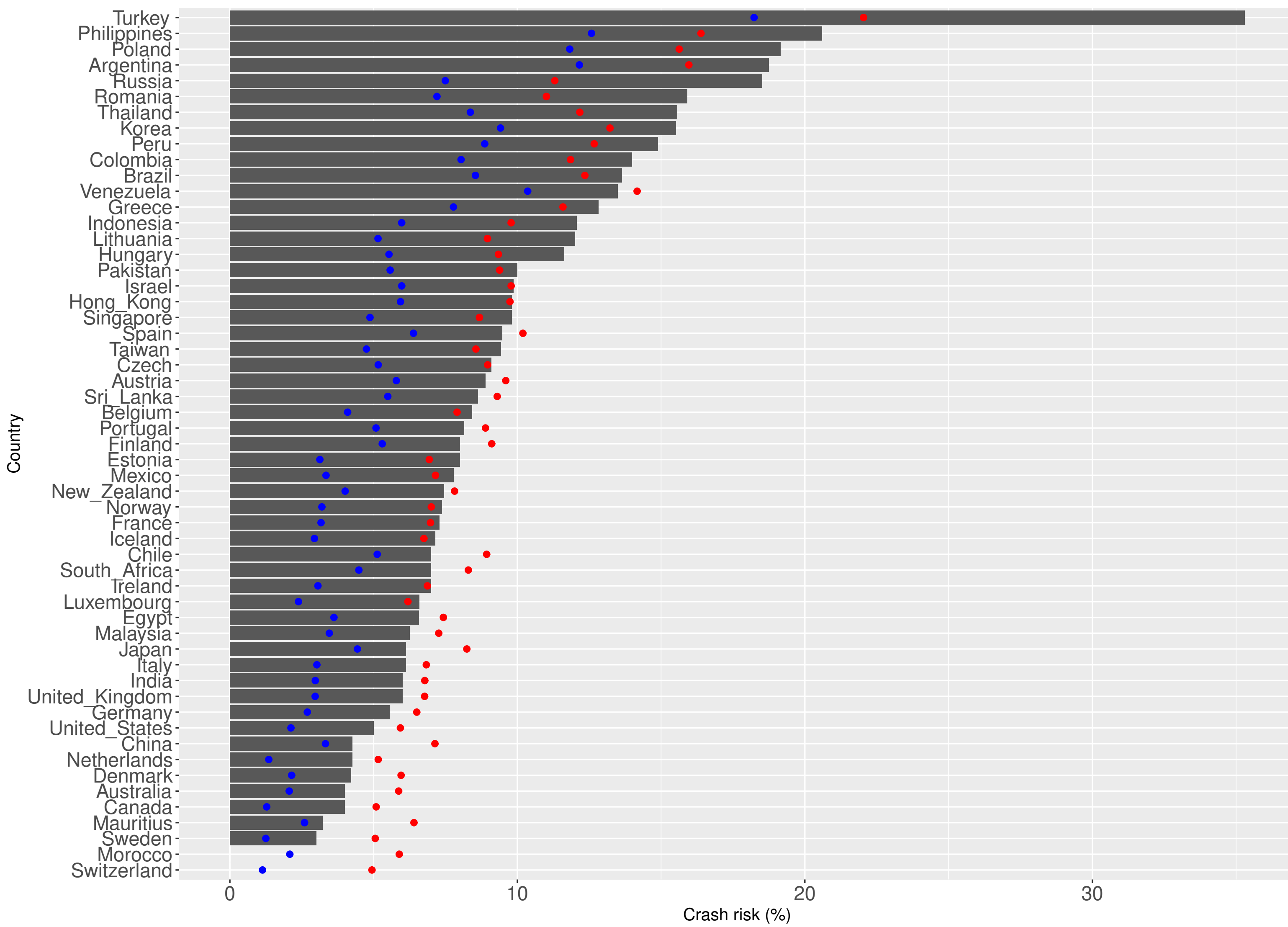


Figure 6. Subjective Crash Risk, Idiosyncratic Crash Risk & In-sample Frequency

The figure plots the posterior mean crash risk (%, red) at the end of the sample (March 2020), and the idiosyncratic crash risk (%, blue) against the historical frequency (%) of market crashes for each country. The posterior estimates are derived from the model in Section II.C.2, with summary statistics in the fourth column in Table III. The red dot denotes the posterior mean crash risk estimate, which equals $p_H(p^S + p_i^I) + p_L p_i^I$. The blue dot denotes the idiosyncratic crash risk p_i^I. The gray bar illustrates the historical frequency of market crashes in the sample.

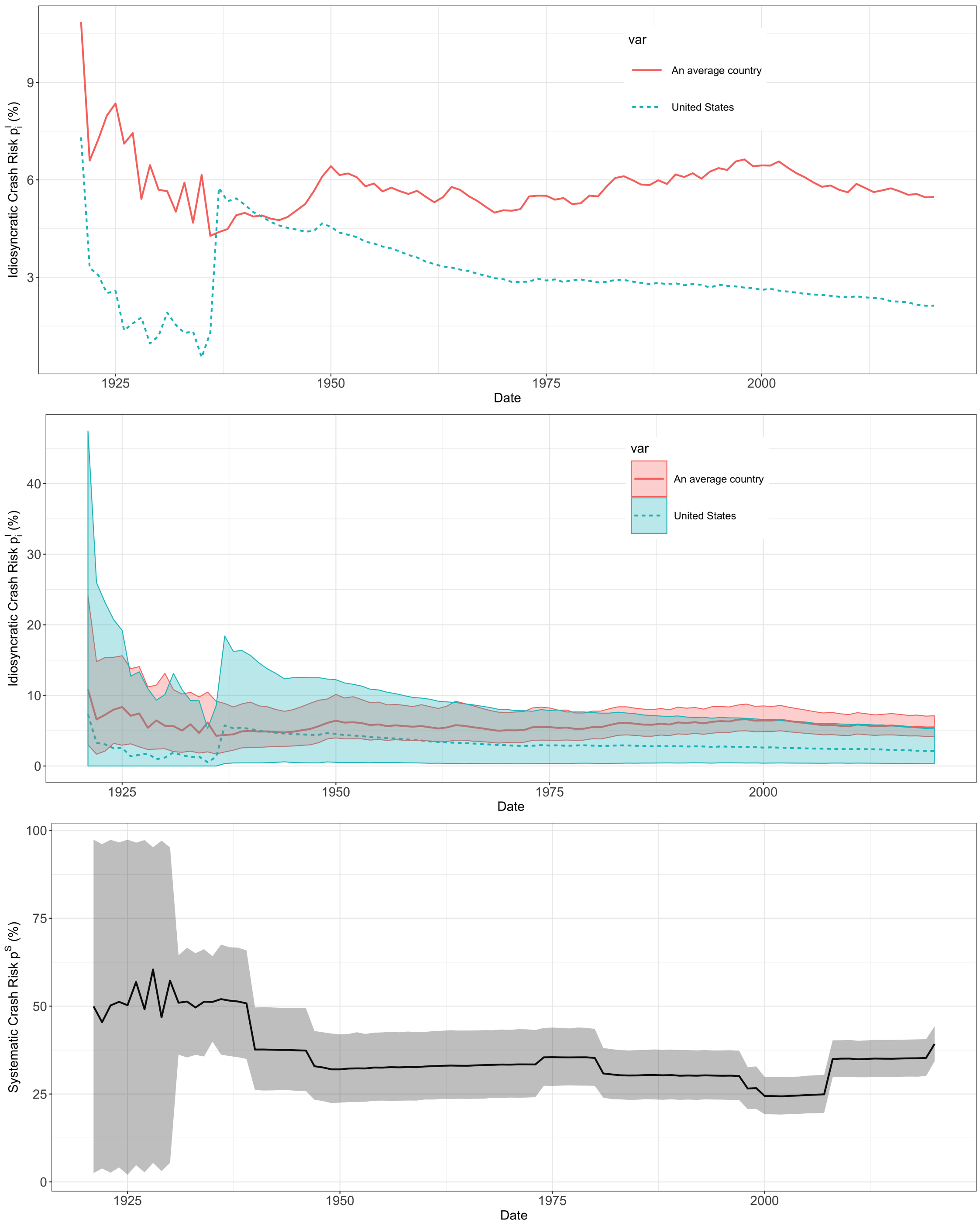


Figure 7. Subjective Crash Risk (Including Systematic Risk) – 1921-2020

The figure shows the time series evolution of posterior mean crash risk over the full sample from the model in Section II.C.2. The upper panel plots the posterior mean idiosyncratic crash risk. The middle 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 US crash risk is shown in the dashed blue line. The lower panel plots the evolution of the systematic risk p^S, and its 95% simulation-based credible set.

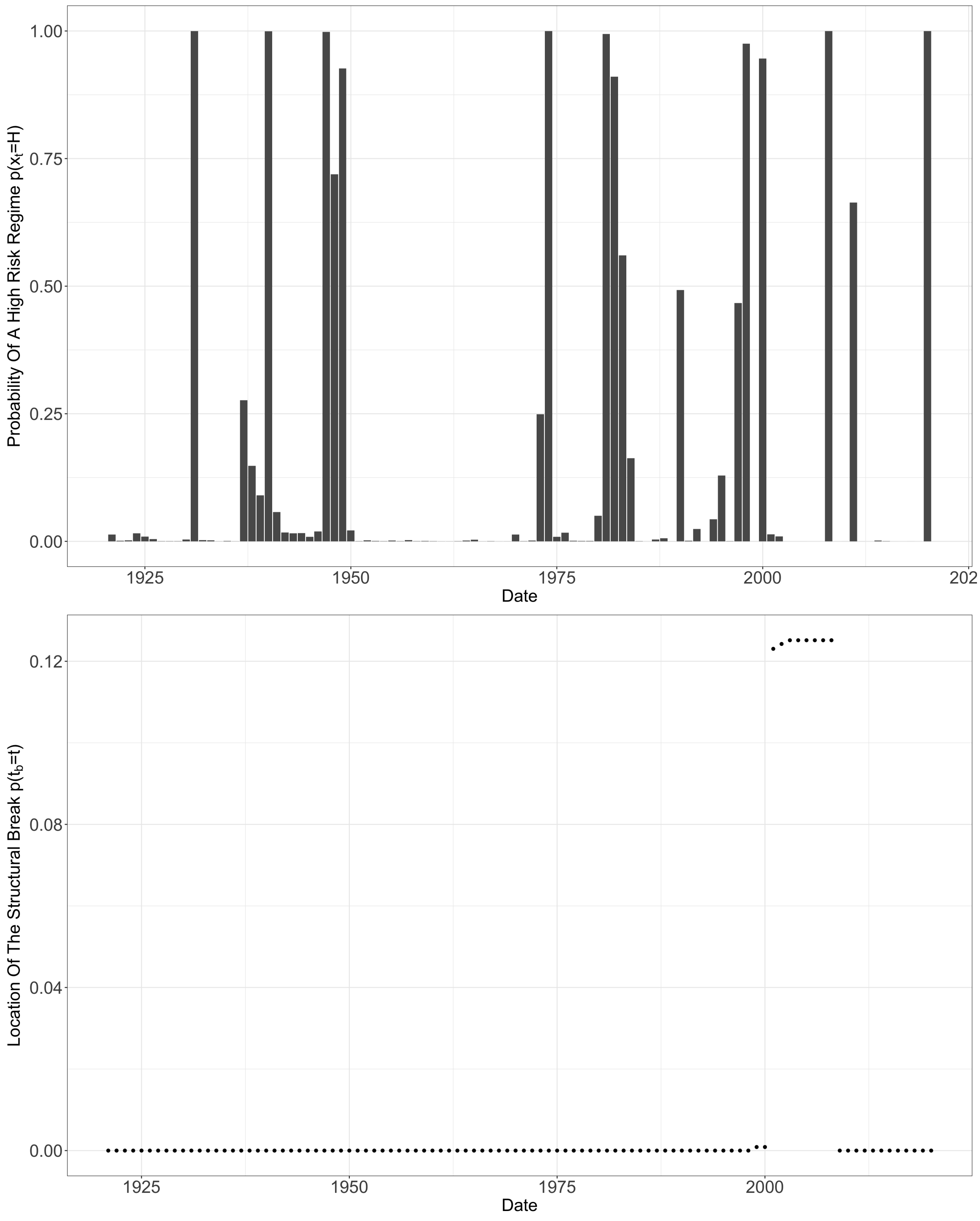


Figure 8. Latent State and the Location of the Structural Break

Based on the filtering algorithm discussed in Section II.C.3, the top panel plots the posterior probability of being in a high risk regime at each point in time, according to Equation 11. The bottom panel shows the probability that a structural break took place (Equation 14), with probabilities summing to 1. Note that the former is a real-time result, which use information only up to time t, whereas the latter uses the full history to infer the location of the structural break.

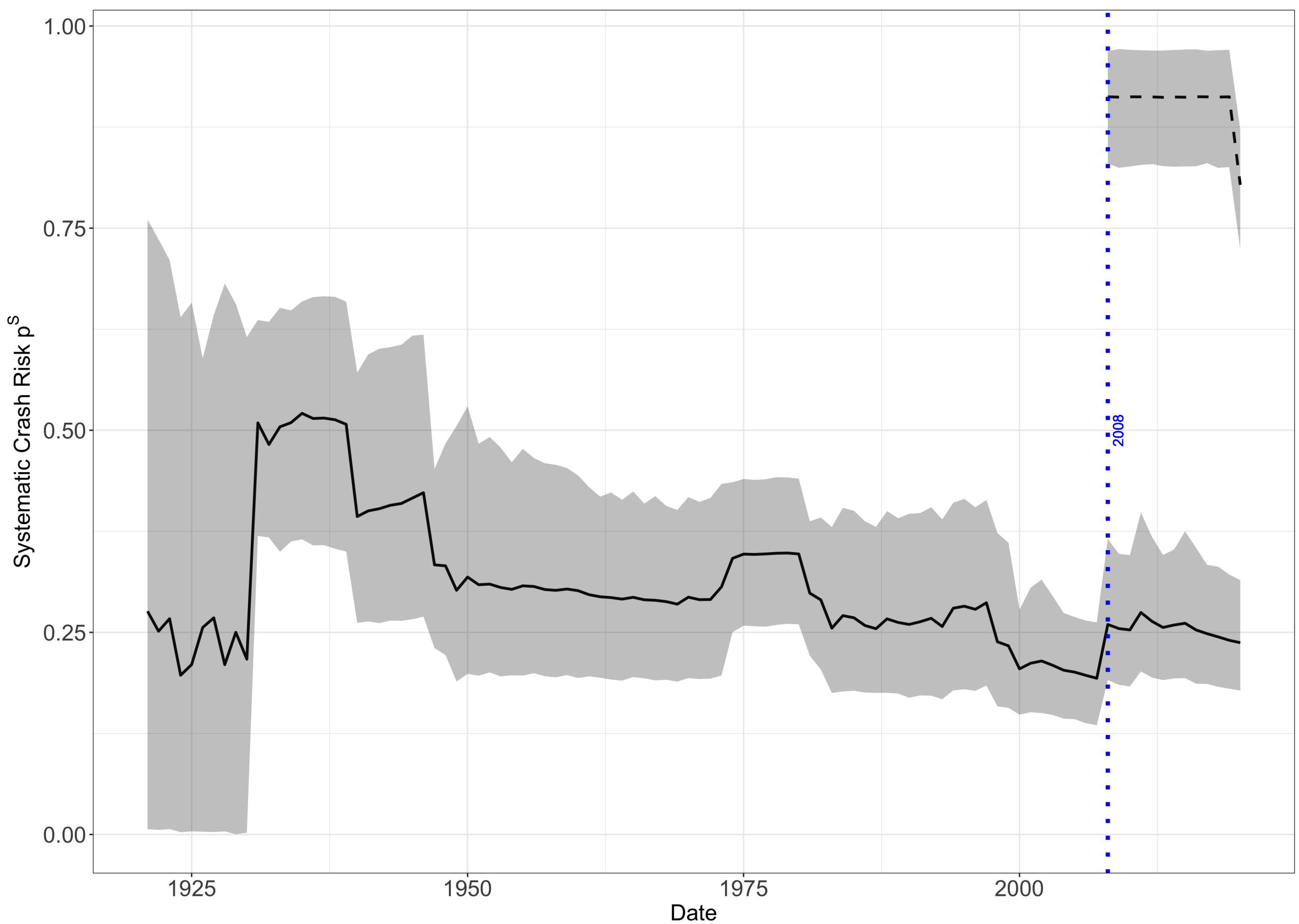


Figure 9. Evolution of Systematic Crash Risk

The figure shows the evolution of systematic risk p^S estimated from the model in Section II.C.3. The solid line traces the posterior mean systematic crash risk before the structural break p_1^S, and the dashed line denotes that after the structural break p_2^S. The dashed blue line indicates where the structural break took place. The shaded areas are the 95% credible sets.

Figure 10. Subjective Crash Risk, Idiosyncratic Crash Risk & In-sample Frequency – 2020

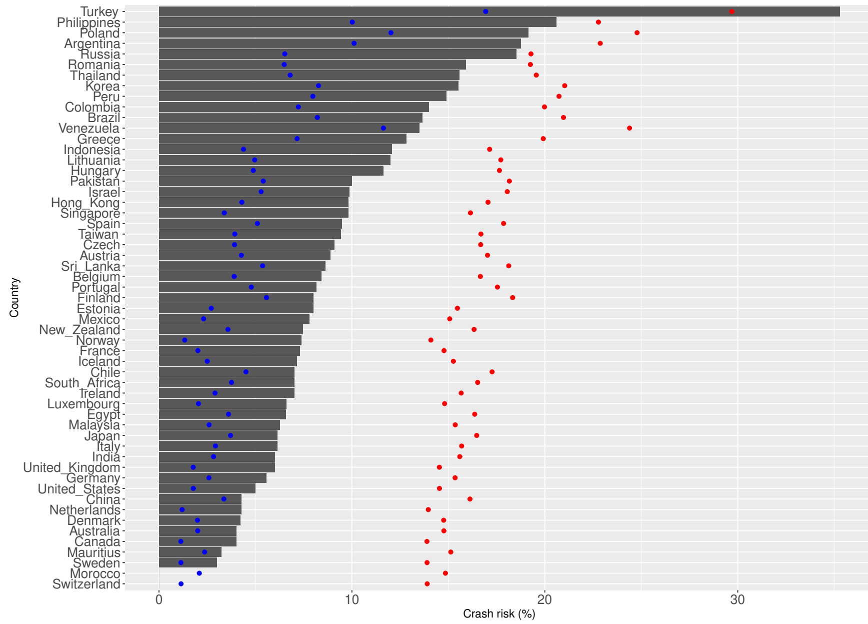
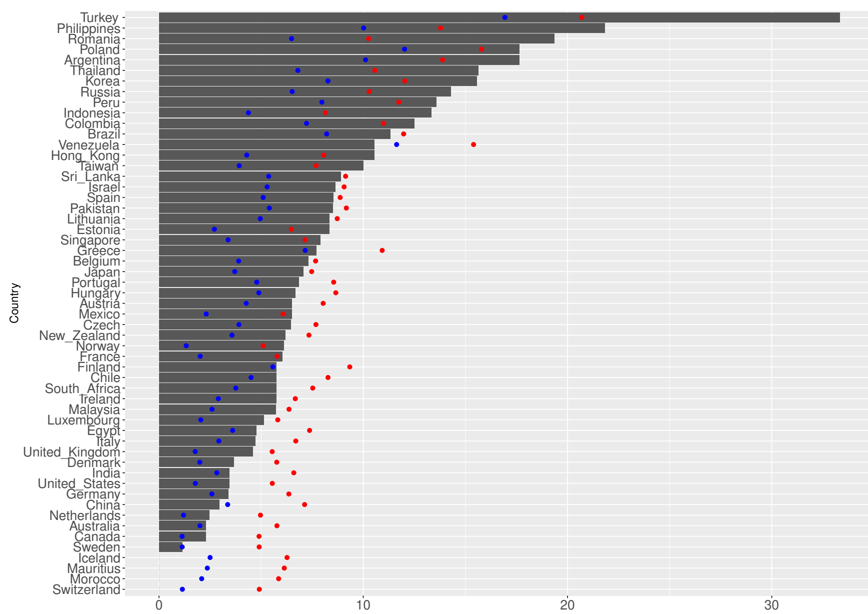


Figure 11. Subjective Crash Risk, Idiosyncratic Crash Risk & In-sample Frequency – 2007



The figures plot the posterior mean crash risk (%, red), and the idiosyncratic crash risk (%, blue) against the historical frequency (%) of market crashes **before (2007) and after (2020)** the structural break. The posterior estimates are derived from the model in Section II.C.3, with summary statistics in the last column in Table III. The red dot denotes the posterior mean crash risk estimate, which equals the unconditional crash risk: $p_H(p^S + p_i^I) + p_L p_i^I$. The blue dot denotes the idiosyncratic crash risk p_i^I. The gray bar illustrates the historical frequency of market crashes in the sample.

Figure 12. Idiosyncratic Crash Risk (Bayesian Filter) – 1921-2020

The figure shows the time series evolution of posterior mean crash risk from the model in Section II.C.3. 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 US crash risk is shown in the dashed blue line.

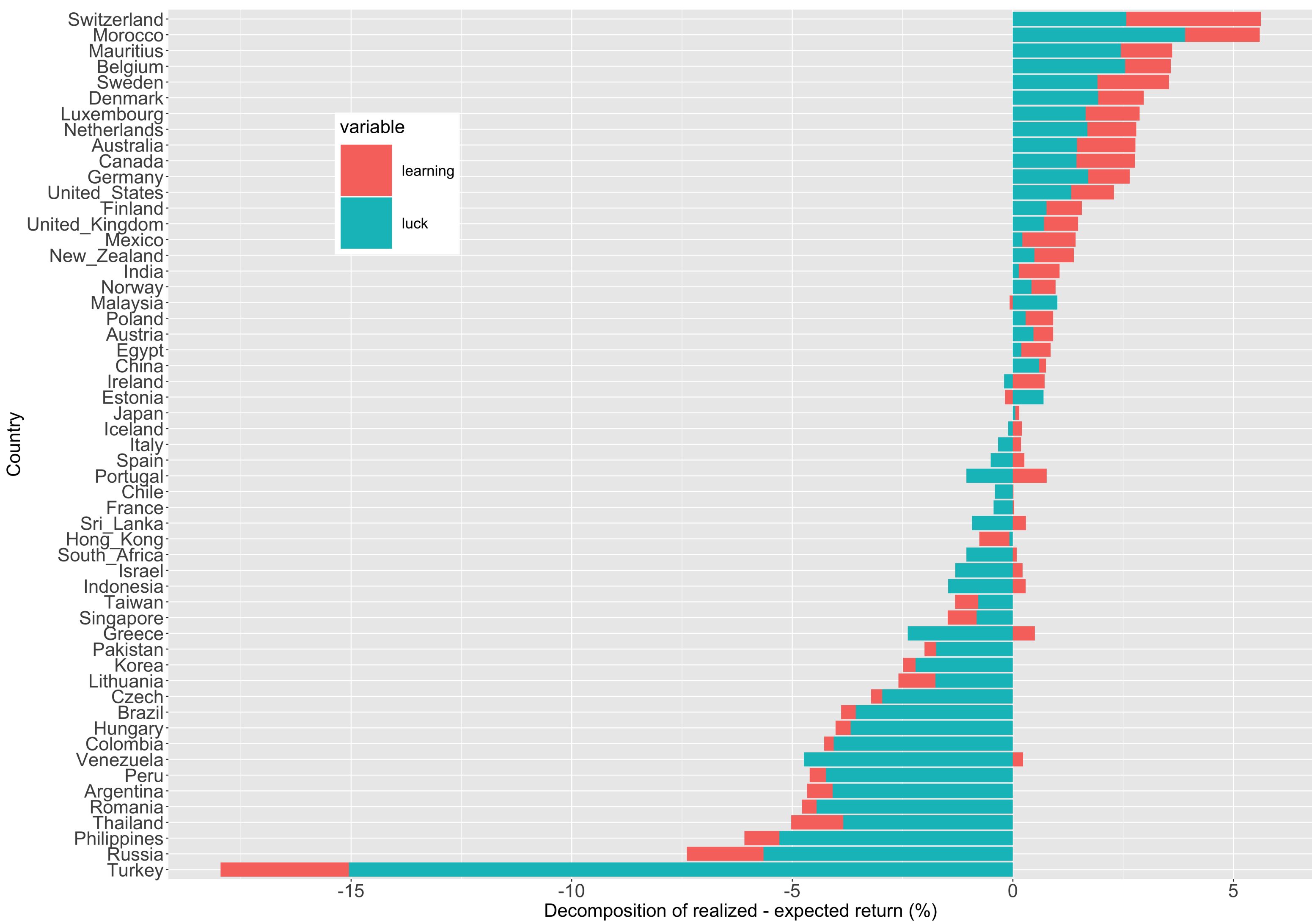


Figure 13. Decomposition Of The Equity Premium

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 as the perceived disaster risk. We simulate a sample of returns of length $L = 100$ for $N = 5000$ times, and compare the average realized return with the expected rate of returns. We use the first $k = 10$ years to initialize the prior and treat them as burn-ins.

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: Posterior Estimates

The table presents the prior and posterior from the three models discussed in the paper. Idiosyncratic Crash Risk refers to the model in Section II.C.1, where crashes are conditionally independent across markets. Systematic Crash Risk corresponds with Section II.C.2 where every country shares a common systematic crash risk component, and systematic risk periods are included in the investor's information set. Filter refers to the model in Section II.C.3, where the investor does not observe the latent states and infers the existence of a structural break in the magnitude of the systematic risk. The posterior estimates are derived with the method discussed in III.C. We present the posterior means and the simulation-based 95% *credible sets in brackets. Note that the credible set is asymmetric by nature.*

Variable	**Prior**	**Idiosyncratic Crash Risk (%)**	**Systematic Crash Risk (%)**	**Filter (%)**
Argentina	Beta(α, β)	14.51	12.16	10.11
		[8.56, 22.02]	[6.16, 19.94]	[4.32, 18.19]
Australia	Beta(α, β)	5.77	2.06	2.01
		[2.56, 9.93]	[0.33, 5.17]	[0.26, 5.2]
Austria	Beta(α, β)	8.98	5.79	4.27
		[4.82, 14.26]	[2.33, 10.72]	[1.07, 9.24]
Belgium	Beta(α, β)	8.68	4.1	3.9
		[4.72, 13.7]	[1.24, 8.5]	[1.03, 8.42]
Brazil	Beta(α, β)	11.7	8.54	8.21
		[6.47, 18.4]	[3.7, 15.24]	[3.37, 15.02]
Canada	Beta(α, β)	5.77	1.28	1.13
		[2.53, 10.02]	[0.06, 3.96]	[0.03, 3.78]
Chile	Beta(α, β)	7.74	5.13	4.51
		[4.02, 12.52]	[2.02, 9.6]	[1.38, 9.1]
China	Beta(α, β)	6.83	3.33	3.36
		[2.59, 12.61]	[0.53, 8.3]	[0.44, 8.77]
Colombia	Beta(α, β)	12.24	8.04	7.22
		[7.39, 18.31]	[3.69, 13.91]	[2.85, 13.16]
Czech	Beta(α, β)	9.16	5.16	3.91

Variable	**Prior**	**Idiosyncratic Crash Risk (%)**	**Systematic Crash Risk (%)**	**Filter (%)**
		[4.2, 15.91]	[1.3, 11.57]	[0.24, 10.48]
Denmark	Beta(α, β)	5.99	2.15	2
		[2.69, 10.36]	[0.33, 5.39]	[0.14, 5.54]
Egypt	Beta(α, β)	7.64	3.62	3.6
		[3.6, 12.87]	[0.88, 8.19]	[0.81, 8.23]
Estonia	Beta(α, β)	8.81	3.13	2.71
		[3.41, 16.37]	[0.18, 9.39]	[0.08, 8.83]
Finland	Beta(α, β)	8.42	5.3	5.57
		[4.53, 13.32]	[2.11, 9.8]	[2.13, 10.57]
France	Beta(α, β)	7.95	3.17	2.01
		[4.07, 12.93]	[0.77, 7.14]	[0.09, 5.64]
Germany	Beta(α, β)	7.07	2.69	2.59
		[3.17, 12.21]	[0.42, 6.83]	[0.36, 6.76]
Greece	Beta(α, β)	11.39	7.78	7.16
		[6.5, 17.64]	[3.35, 13.92]	[2.66, 13.43]
Hong Kong	Beta(α, β)	9.51	5.94	4.29
		[4.58, 16.1]	[1.82, 12.18]	[0.51, 10.61]
Hungary	Beta(α, β)	10.31	5.54	4.89
		[4.97, 17.27]	[1.38, 12.38]	[0.91, 11.33]
Iceland	Beta(α, β)	8.45	2.94	2.5
		[3.29, 15.59]	[0.17, 8.96]	[0.07, 8.29]
India	Beta(α, β)	7.1	2.97	2.83
		[3.56, 11.56]	[0.71, 6.64]	[0.57, 6.56]
Indonesia	Beta(α, β)	10.73	5.98	4.38
		[5.69, 17.32]	[1.85, 12.16]	[0.67, 10.85]
Ireland	Beta(α, β)	7.74	3.06	2.9
		[4, 12.42]	[0.75, 6.89]	[0.65, 6.7]
Israel	Beta(α, β)	9.58	5.98	5.29
		[5.04, 15.48]	[2.11, 11.7]	[1.52, 11.05]
Italy	Beta(α, β)	7.18	3.03	2.93

Variable	**Prior**	**Idiosyncratic Crash Risk (%)**	**Systematic Crash Risk (%)**	**Filter (%)**
		[3.57, 11.8]	[0.71, 6.84]	[0.64, 6.72]
Japan	Beta(α, β)	7.19	4.44	3.71
		[3.58, 11.83]	[1.58, 8.66]	[0.84, 8.25]
Korea	Beta(α, β)	12.54	9.41	8.27
		[7.06, 19.65]	[4.14, 16.67]	[2.81, 16.22]
Lithuania	Beta(α, β)	10.16	5.15	4.96
		[4.44, 18.22]	[0.87, 12.76]	[0.67, 13.43]
Luxembourg	Beta(α, β)	7.55	2.39	2.05
		[3.79, 12.37]	[0.38, 6]	[0.15, 5.66]
Malaysia	Beta(α, β)	7.79	3.46	2.6
		[3.27, 13.82]	[0.55, 8.61]	[0.08, 8.08]
Mauritius	Beta(α, β)	6.92	2.6	2.37
		[2.31, 13.39]	[0.14, 7.97]	[0.07, 7.91]
Mexico	Beta(α, β)	8.29	3.34	2.31
		[4.32, 13.37]	[0.79, 7.53]	[0.25, 6.02]
Morocco	Beta(α, β)	5.09	2.09	2.09
		[1.31, 10.4]	[0.11, 6.42]	[0.07, 6.9]
Netherlands	Beta(α, β)	6.01	1.35	1.2
		[2.61, 10.4]	[0.06, 4.24]	[0.03, 4.07]
New Zealand	Beta(α, β)	8.07	4.01	3.57
		[4.17, 13.01]	[1.21, 8.29]	[0.87, 7.91]
Norway	Beta(α, β)	8.02	3.2	1.33
		[4.16, 12.92]	[0.8, 7.1]	[0.03, 4.55]
Pakistan	Beta(α, β)	9.64	5.58	5.4
		[4.86, 15.9]	[1.75, 11.42]	[1.36, 11.85]
Peru	Beta(α, β)	12.87	8.87	7.98
		[7.97, 18.89]	[4.29, 14.93]	[3.58, 14.11]
Philippines	Beta(α, β)	15.68	12.58	10.02
		[9.71, 23.22]	[6.4, 20.66]	[3.84, 18.42]
Poland	Beta(α, β)	13.98	11.82	12.02

Variable	**Prior**	**Idiosyncratic Crash Risk (%)**	**Systematic Crash Risk (%)**	**Filter (%)**
		[7.82, 21.99]	[5.51, 20.46]	[5.17, 21.41]
Portugal	Beta(α, β)	8.53	5.08	4.78
		[4.44, 13.73]	[1.8, 9.88]	[1.44, 9.78]
Romania	Beta(α, β)	12.34	7.2	6.49
		[6.48, 20.11]	[2.22, 14.87]	[1.58, 13.92]
Russia	Beta(α, β)	12.48	7.49	6.52
		[6.13, 21.21]	[1.82, 16.79]	[1.33, 15.31]
Singapore	Beta(α, β)	9.5	4.87	3.38
		[4.6, 16.02]	[1.23, 10.83]	[0.14, 9.32]
South Africa	Beta(α, β)	7.75	4.49	3.76
		[4.01, 12.5]	[1.57, 8.76]	[0.93, 8.21]
Spain	Beta(α, β)	9.37	6.39	5.1
		[5.23, 14.62]	[2.75, 11.36]	[1.65, 10.17]
Sri Lanka	Beta(α, β)	8.89	5.49	5.37
		[4.3, 14.9]	[1.69, 11.31]	[1.44, 11.55]
Sweden	Beta(α, β)	5.12	1.25	1.14
		[2.07, 9.19]	[0.06, 3.91]	[0.03, 3.9]
Switzerland	Beta(α, β)	3.15	1.14	1.14
		[0.72, 6.76]	[0.06, 3.57]	[0.03, 3.83]
Taiwan	Beta(α, β)	9.31	4.75	3.93
		[4.46, 15.69]	[1.18, 10.56]	[0.35, 10.17]
Thailand	Beta(α, β)	12.18	8.37	6.8
		[6.46, 19.78]	[3.03, 16.17]	[1.64, 14.87]
Turkey	Beta(α, β)	19.73	18.23	16.93
		[11.25, 30.89]	[8.72, 30.9]	[7.79, 29.52]
United Kingdom	Beta(α, β)	7.1	2.97	1.77
		[3.53, 11.65]	[0.72, 6.61]	[0.07, 5.2]
United States	Beta(α, β)	6.43	2.12	1.78
		[3.02, 10.83]	[0.33, 5.35]	[0.07, 5.2]
Venezuela	Beta(α, β)	11.92	10.36	11.63

Variable	Prior	Idiosyncratic Crash Risk (%)	Systematic Crash Risk (%)	Filter (%)
		[7.11, 17.96]	[5.6, 16.5]	[5.98, 18.7]
$\alpha/(\alpha+\beta)^1$	$(\alpha+\beta)^{-5/2}$	9.25	5.47	4.86
		[7.88, 10.84]	[4.2, 7.11]	[3.46, 6.7]
$\alpha+\beta^2$	$(\alpha+\beta)^{-5/2}$	5509.16	2911.73	2756.77
		[2528.7, 11370.45]	[1408.13, 5452.55]	[1209.85, 5436.72]
${p_{HL}}^3$	Beta(1,1)		90.87	66.69
		-	[69.24, 99.74]	[37.21, 96.39]
${p_{LH}}^4$	Beta(1,1)		9.76	12.59
		-	[4.58, 16.55]	[6.11, 21.13]
${p_S}^5$	Uniform(0, 1)		39.26	
		-	[34.33, 44.29]	-
${p_S^1}^6$	Uniform(0, 1)			23.73
		-	-	[17.81, 31.46]
${p_S^2}^7$	Uniform(0, 1)			80.36
		-	-	[72.38, 87.19]
${p_H^0}^8$	Uniform(0, 1)			47.56
		-	-	[2.28, 97.08]

[1] Global mean idiosyncratic crash risk
[2] Effective sample size of the global crash risk distribution (smaller values mean more dispersion)
[3] Probability of transitioning from a high-risk state to a low-risk state
[4] Probability of transitioning from a low-risk state to a high-risk state
[5] Systematic risk
[6] Systematic risk pre structural break
[7] Systematic risk post structural break
[8] Probability of starting off in a high-risk state in 1920