Electronic copy available at: http://ssrn.com/abstract=1977867

1 Introduction

Why are the stock returns so high and so volatile? This is a classic question in both economics and finance. This paper attempts to answer the above question by developing a consumption based asset pricing model in which agents learn about rare events that affect their dividend stream.

A rare disaster is defined as some infrequently occurring event with long-lasting effect on consumption growth based on the influential work by Rietz (1988) and Barro (2006). While on average rare disasters have a negative long-run effect on consumption, it is important to note that there is a small possibility that a disaster can have a positive long-run effect.[1] Observations of historical events underline the importance of the uncertainty about (potential) disasters and provide the key motivation for this paper. During the 2007-2009 financial crisis, many commentators, including well-known macroeconomists, have highlighted the possibility that the U.S. economy could fall into another Great Depression and the markets reacted with a large drop in stock prices and an increase in volatility. Another recent example is the nuclear facility accident in Japan in early 2011 which resulted in a 22% drop of the Japanese stock market within two days of the accident, as the country faced uncertainty over the question wether they would be able to contain the accident or wether it would develop into a nuclear disaster. Learning about rare events is an important driver for asset price volatility in our model as the uncertainty about future consumption created by rare events can lead to large movements in asset prices.

The majority of existing work on rare disaster models generally assumes that the entire damage caused by a rare disaster occurs in a single time period, and that agents have complete information about the disaster. We model disasters in a more realistic way, following Barro, Nakamura, Steinsson and Ursua (2011), that the disaster unfolds over multiple periods. We also extend Barro et al. (2011) to allow for two sources of uncertainty about a rare disaster: (1) whether a disaster has occurred is not directly observable since its effect takes time to unfold; (2) how much damage a disaster will cause is governed by some unknown parameters due to the lack of historical data for such a rare event. The evidences for these two types of uncertainty are apparent in the empirical

[1] A long run positive impact of a disaster could for example be due to structural changes that are resulting from a disaster through behavioral changes/political reforms.

Electronic copy available at: http://ssrn.com/abstract=1977867

results of Barro et al. (2011). The estimated posterior probability of being in a disaster vary significantly as illustrated in Figure IV of their paper. The estimated standard deviations of the short-run and the long-run shocks during disasters are large, revealing that there is a great amount of uncertainty during a disaster about its short-run and long-run effects. The authors find that the short-run effect on average is about twice as large as the long-run effect, indicating that disasters are followed by some form of recovery.

Agents in our model observe real-time data on consumption and update their beliefs about the occurrence and the severity of a disaster over time. Learning implies an endogenously time-varying risk premium because the time variation of beliefs generates time variation in how risky the current economy is as perceived by agents. Since it is not directly observable whether a disaster occurs, agents can sometimes mistake a recession for the beginning of a rare disaster. Misperceived risk su ers from one of the following two problems. Many models including Barro (2006) and Barro et al. (2011) imply a constant dividend-price ratio during non-disaster periods. This makes it impossible for the dividend-price ratio to predict time varying returns since disasters are rare events and most of the time the economy is not in a disaster state. Another stream of models such as Gabaix (2008) and Gourio (2011) feature time-varying dividend-price ratios during non-disaster periods purely due to the exogenous time variation in risk. This feature again deprives the dividend-price ratio of its predictability for future returns in the absence of exogenous jumps in disaster risk. In addition, what remains unexplained in these papers is the fundamental driving force behind the time-varying disaster probability. In this paper, learning is explicitly modeled so that agents' learning about the occurrence and the severity of a rare disaster endogenously generates time-variation in the perceived risk of a disaster and, in turn, in both dividend-price ratios and equity returns. In this sense, our model provides a foundation for why the disaster probability can be assumed to be time-varying. At the same time, it also contributes to understanding the predictability in the data (see also Brandt et al.,2004 and Cogley and Sargent, 2008).

Second, we can use our model to interpret the historical consumption data of the 20th century and can compare realized asset returns with model implied asset returns following Campbell and Cochrane (1999). This supports a better understanding of the key model mechanism. Cyclical variations in consumption growth will induce agents in the model to learn and update their beliefs about disasters, which will in turn imply a time series of equity returns. We can then compare historical data on equity returns with the model-implied ones, as a test for the performance of our model.

1.1 Literature Review

It is well-known that there is a long list of stock market and bond market puzzles in macro- nance literature. Among them, time-varying risk premium havetime69a 0 -22(e0 -22ompta 0 -22m316(e)-

where the most successful work is Campbell and Cochrane (1999). With habit being a weighted varying risk premium, e.g. Veronesi (1999, 2004), Chen and Pakos (2007). However, focusing on only one source of uncertainty limits the explanation power of a model. Without state uncertainty, parameter uncertainty will die out in the long run as more data is accumulated and agents are increasingly confident about their estimates for the constant parameters. Without parameter uncertainty, learning the hidden state tends to be too fast to play an important role in explaining time-varying risk premium, given the limited variability of observed consumption, output or dividend process, which typically serves as a noisy signal for the hidden state.

A unified framework to study jointly parameter and state uncertainty is thus promising in bringing the learning models closer to the data. However, very little work has been done so far on this front. One pioneer paper is Lewis (1989), which shows that a simple learning model in the presence of these two sources of uncertainty is capable to explain the behavior of U.S. dollar–German mark forecast errors during the early 1980's. This paper pushes further in developing the unified learning framework given that learning about a rare disaster is a natural combination of parameter and state uncertainty: 1) the lack of ex-ante knowledge of a rare disaster forces agents to learn its unknown parameters; and 2) the occurrence of a rare disaster is unobservable directly and is in turn a hidden state.

In modeling the parameter uncertainty, we are taking a consistent approach between the learning and mapping into asset prices. In other words, our agents are rational learners so they takes into account the parameter uncertainty in pricing assets. Although this idea of rational learning is widely adopted in the learning literature,[3] this paper is the first - to the best of our knowledge - to implement it in a model with both parameter uncertainty and state uncertainty. In a related study by Johannes, Lochstoer and Mou (2010) where the two-sided uncertainties are also present, agents are assumed to ignore the parameter uncertainty in pricing of assets. To investigate the impact of this simplification assumption, we also compute asset returns under this assumption and contrast them with the ones computed using the fully rational approach. We find that ignoring the parameter uncertainty in asset pricing significantly overstates the volatility of returns.

[3] Townsend (1978), Wieland (2000a, 2000b), Guidolin and Timmermann (2007), Cogley and Sargent (2008).

The two-sided uncertainty distinguishes our model from the existing literature on models with learning, where typically only one kind of uncertainty is present. With perfect knowledge of $_t$ and only uncertainty about I_t, the model reduces to a standard hidden Markov regime switching model. With perfect knowledge of the disaster state I_t and only uncertainty about the disaster severity $_t$, the model ts into the familiar framework of adaptive learning. $_t$

Conditional on $_t$ and I_t; the likelihood function for consumption growth c_t is the density function for a normal random variable. The likelihood function for c_t can be combined with Bayes' rule to update agents' belief about the disaster state, I_t, (the disaster parameter, $_t$) conditional on perfect knowledge of the disaster parameter, $_t$ (the disaster state I_t).. However, if agents have to learn both { state and parameter { the simple Bayes' updating rule does not apply any longer. The next section explains in detail how agents update their belief in a world

2.2.1 Learning the State under Parameter Uncertainty

Without perfect knowledge of $_t$, the posterior belief of I_t, $\Pr(I_t = 1|c^t)$, can still be obtained by integrating out $_t$ from its conditional information set, as long as agents have a belief about $_t$. The resulting likelihood function of c_t is independent of a particular $_t$ consequently becomes

$$\Pr\ c_t|I_{t-1} = i; I_t = j; c \qquad {}_t j\ t i; \mathrm{t}$$

$$\Pr\ {}_t|I_{t\ i;I}\ \mathrm{ti;\ I}$$

2.2.2 The Learning Switch

Each time a disaster starts, the parameter $_t$ is drawn from F ($_t$)and will remain constant until the end of this particular disaster. Thus, agents' belief about $_t$ has to be conditioned on the state of the economy. However, since agents in our model never know the disaster state perfectly, the number of possible disasters will grow with the number of periods, so does the number of possible $_t$s that agents have to form belief about. The exploding number of beliefs about di erent $_t$s over time make this problem intractable. To resolve this issue, we put more structure on how agents update their beliefs and introduce an instrument called "learning switch". That is, agents

Therefore, the implication of the learning switch being o at period t is that there are only

When the learning switch is o at period t, it can be turned on by evidence in period t, if

$$\frac{\Pr(I_{t-1} = 0; I_t = 1 | c^t)}{\Pr(I_{t-1} = 0; I_t = 0 | c^t)} > T_{on} \tag{2.4}$$

The utility function of agents is Epstein-Zin (EZ) with γ being the coefficient of relative risk aversion and ψ being the inter-temporal elasticity of substitution (IES). The time discount factor is β.

3.1 Asset Pricing with Perfect Information

Under EZ utility, if agents can observe μ_t and I_t perfectly, we know that the price-dividend ratio (henceforth called PDR), P_t/C_t, is a function $f(\mu_t; I_t)$ – PDR function – which satisfies the following Euler equation.

$$
\begin{aligned}
(P_t/C_t)^{\theta} &= E_t\left[\beta^{\theta}(C_{t+1}/C_t)^{\frac{\theta(\psi-1)}{\psi}}(P_{t+1}/C_{t+1}+1)^{\theta}\right] \text{ so that:} \\
f(\mu_t; I_t) &= \left(E_t\left[\beta^{\theta}\exp\left(\frac{\theta(\psi-1)}{\psi}c_{t+1}\right)[f(\mu_{t+1}; I_{t+1})+1]^{\theta}\right]\right)^{1/\theta} \qquad (3.2)
\end{aligned}
$$

where $\theta = (1-\gamma)(1-1/\psi)^{-1}$.

After obtaining the price-to-dividend ratio, the returns of two assets are simply:

$$
\begin{aligned}
R^f_{t+1}(\mu_t; I_t) &= \left[E_t\,\beta^{\theta}\exp\left(\left(\frac{\theta(\psi-1)}{\psi}-1\right)c_{t+1}\right)\left(\frac{f(\mu_{t+1}; I_{t+1})+1}{f(\mu_t; I_t)}\right)^{\theta-1}\right]^{-1} \\
R^e_{t+1}(\mu_{t+1}; I_{t+1}) &= \frac{f(\mu_{t+1}; I_{t+1})+1}{f(\mu_t; I_t)}\exp(c_{t+1})
\end{aligned}
$$

3.2 Asset Pricing with Imperfect Information

However, neither μ_t nor I_t is directly observable, so we have to replace them in the PDR function $f(\mu_t; I_t)$ by the corresponding agents' beliefs, $\Pr(\mu_t|I_t=0; c^t)$, $\Pr(\mu_t|I_t=1; c^t)$ and $\Pr(I_t|c^t)$. $\Pr(\mu_t|I_t; c^t)$ is normal distribution and thus can be summarized by its mean and variance. $\pi_t = \Pr(I_t|c^t)$ is a scalar variable. The PDR function under imperfect information, $F[\Pr(\mu_t|I_t; c^t); \Pr(I_t|c^t)]$, needs to satisfy:

$$F\left(\Pr\left(_{t}|I_t;c^t\right);\ _t\right) = \left(E_t\left[\exp\left(\frac{(\ \ 1)}{}c_{t+1}\right)\left\{F\left(\Pr\left(_{t+1}|I_{t+1};c^{t+1}\right);\ _{t+1}\right)+1\right\}\right]\right)^{1=}:$$

In addition to replacing $(_t; I_t)$ by the beliefs, imperfect information about $_t$ and I_t also changes the expectation formation E_t. In particular, conditional on the information set at period t, the distribution of future $c_{t+1} = \ + I_{t+1}\ _{t+1} + \ _{t+1}$ is determined by $\Pr(_t|I_t; c^t)$ and $\Pr(I_t = 1|c^t)$ due to the persistence of $_t$ and I_t. Each future data c_{t+1} implies a set of updated state variables, $\Pr(_{t+1}|I_{t+1}; c^t; c_{t+1})$ and $\Pr(I_{t+1} = 1|c^t; c_{t+1})$, which are in turn associated with a particular future PDR. The expectation is taken by averaging across all possible future c_{t+1} and the associated PDRs. Details for the computation of the price-dividend ratio can be found in the Appendix. A similar methodology applies to the computation of the risk free rate.

To the best of our knowledge, this is the first paper that treats the belief about the parameter - a distribution - as the state variable in the PDR function. A typical approach in existing asset pricing models with parameter uncertainty is to assume that agents are adaptive learners in the sense that they view parameters as constants in forming expectations whereas their beliefs about parameters are in fact updated over time when new data comes in [e.g. Johannes et al. (2010)]. Our model setup allows us to overcome this inconsistency between expectation formation and learning so our agents are fully Bayesian. This new approach thus takes full account of how parameter uncertainty affects asset returns.

In order to better understand the implications of this new approach on asset returns, we also compute asset returns using two alternative approaches with restrictive assumptions imposed at various steps during the computation. The first approach is the common one taken in the existing literature, where is fixed at the level of its posterior mean during the PDR computation. The second approach uses the posterior belief about in determining the distribution of future c_{t+1} but does not update the posterior belief about with each c_{t+1} in evaluating the next-period price-dividend ratio.. We view the second approach as a step in between the first approach and the benchmark one since this approach does take into account the belief about but treats it as xed rather than a state variable.

Another novelty of our model is the presence of the dual uncertainties, namely parameter uncertainty and state uncertainty, in agents' learning. To gain a better understanding of its impact, we compute asset returns under three information structures: The rst scenario is no learning where agents have perfect information about the current and past states of the economy $(I_s)_{s=0}^{t}$. We rst assume that agents also have perfect information about the future s so that there is no uncertainty in the model. Then we assume agents view the parameter as a new draw from the distribution F ( ) each period. This case shuts down the learning and is designed to capture the e ect of parameter uncertainty on asset returns. The second scenario is partial learning where agents learn about the disaster state I_t conditional on the knowledge of .[5] This case shows the e ect of state uncertainty on asset returns and is one of the well-known hidden Markov regime switching models. The third scenario is our benchmark case { joint learning { where agents have to learn jointly the parameter and the state I_t. Parameter uncertainty and state uncertainty interact with each other in this case and the interaction has important implications for asset returns.

3.3 Calibration

All our results are computed with the same parameter values as reported in Table 1. We set the values of the parameters governing consumption process equal to the posterior means estimated in Barro et al. (2011) whenever possible. The time period is one year. Because the equity return is a leveraged claim on the consumption stream in reality, we compute levered equity return with the leverage parameter set to 2.[6] The preference parameters are standard with a rather low risk aversion coe cient of 6.5, the IES equal to 2.[7] Risk aversion is picked to match the equity return

[5] Another case of partial learning is to let agents learn the parameter conditional on their perfect knowledge of the state I_t. But this case is equivalent to the no learning case with a time-varying F ( ), so we omit it in our computation.

[6] T615 Td 11e61(is)-345(ass/27(andt0 Ton)Td 54(alu38.216 arn)]Ta(y)-mFr)-337ut352(with)-3510the Td [(Alf-3510the)-307

in the benchmark case. An extensive debate evolves around the values of the IES. It is crucial for our model - as for most successful asset pricing models - to have an IES larger than one. We set the time discount factor, , equal to 0.978. The discount factor is picked to roughly match the risk free rate in our benchmark model. The only two free parameters in our learning model are two thresholds governing the learning switch. T_{on} = 0:015 means that the learning switch will be turned on when the probability of a new disaster today is 1.5% of the probability of no disaster today. This implies our agents are quite alerted by the possibility of a new disaster. T_{off} = 50 means that the probability of no disaster today needs to be 50 times larger than the probability of a disaster continuing in order to turn the learning switch o . This implies a dominant evidence is required to assure our agents that there is no need to worry about being in a disaster today. Given the relative arbitrary T_{on} and T_{off}, we compute our joint learning model under alternative choices of T_{on} and T_{off}. The asset returns are of course sensitive to the choices since the number of periods when agents are learning varies, but all the major results are robust within a reasonable range of T_{on} and T_{off}. Results for robustness checks can be found in the Appendix.

[Table 1 about here.]

4 Results in an endowment economy

In this section, we report asset returns in the models with no learning, partial learning and joint learning, respectively. In the no learning model, we rst compute asset returns with "no uncertainty" about either the parameter or the state, and then we compute asset returns with only parameter uncertainty present. In contrast, the partial learning model only has state uncertainty. In the joint learning model, we compute the PDR function following three di erent approaches: 1) our benchmark approach ("dist_state") in which the distribution of is treated as a state variable; 2) an intermediate approach (" x_dist") in which the distribution of is xed to agents' posterior belief; and 3) the common approach (" x_mean") in the literature that xes the parameter at its posterior mean.

model implies a time varying local risk aversion coe cient larger than 30 in simulations.

We further group those results into three sets. The first set of results shows the behavior of equity returns and risk free rates when a disaster occurs. The second set of results reports the statistics of asset prices in a sample without any realization of disasters. The third set of results looks at how our model fits the historical consumption and stock return data.

4.1 Returns over a Sample Disaster

In this subsection, we investigate the dynamics of asset returns with a sample disaster starting at period 5 and lasting for 6 periods (years). The long-run damage on consumption growth of this particular disaster is set to be 4% each period during the disaster, which implies a total 24% drop of consumption in the long run. All the shocks ϵ_t are set to zero.

[Figure 1 about here.]

The left panels of Figure 1 and Figure 2 display the asset returns computed in the aforementioned 6 cases under EZ utility. The black solid line is the time series of equity returns and the black dashed line is the time series of risk free rates. The green line reflects the de-trended log CTf 5.775 -1.

exposure to the parameter uncertainty is much larger than in the normal state. It follows that the equity price plunges deeper at the onset of a disaster when the parameter uncertainty is present.

Now we turn to the partial learning case as displayed in Figure 1. Compared to the case with no uncertainty, both the crash and the boom become more gradual. In other words, the movements in equity returns are smaller but are more persistent. The reason for this change is due to agents' learning about the state I_t. The right panel plots the time series of agents' posterior belief of $I_t = 1$. At the onset of the disaster, agents are not so sure that it is the beginning of a disaster and the posterior belief is as low as less than 20%. This belief quickly peaks up after consecutive observations of low consumption growth and approaches 1 at the end of the disaster. Due to the presence of state uncertainty in the beginning of the disaster, the equity price does not drop as dramatic as in the case with no uncertainty, resulting in a gradual fall of equity return. After the disaster ends at period 10, it takes the agents a couple of periods to be sure that the economy is back to normal state. The equity price thus rises gradually, so does the equity return.

[Figure 2 about here.]

The three cases of joint learning displayed in Figure 2 share the same dynamics of agents' belief of being in a disaster, so the only di erence among them is the way that agents compute asset prices as discussed in section 3. Similar to the belief in the partial learning case, it takes time for agents to learn the occurrence and the ending of a disaster. However, signi cant di erence exists between the partial learning and the joint learning case. As plotted in Figure 3, the di erence is rst positive and then negative, re ecting the fact that agents learn slower in the joint learning case because of the presence of parameter uncertainty. Notice also that the di erence is much larger at the beginning of the disaster than after the disaster ends. This observation con rms the role played by the parameter uncertainty in agents' learning about the state. The onset of the disaster is when the parameter uncertainty is most pronounced so the di erence in learning is larger. As more observations of consumption growth are accumulated over time, the parameter uncertainty shrinks dramatically so the di erence in learning becomes much less signi cant.

[Figure 3 about here.]

By comparing the partial learning case with our benchmark joint learning case where agents are assumed to be fully rational ("dist_state"), it helps us to understand the impact of parameter uncertainty on asset returns when it interacts with the state uncertainty.

The rst feature is that the risk free rates are much lower in the joint learning case. It is due to the higher demand of risk free asset when the risky asset { equity { is exposed to extra risk of parameter uncertainty.

The second feature is the similarity of the size of stock market crash between two cases at the onset of the disaster. It is due to two o setting e ects of the parameter uncertainty. On the one hand, the parameter uncertainty slows agents' learning about the disaster state, which makes agents less pessimistic and in turn mitigates the drop of their demand for equity. On the other hand, the additional parameter uncertainty makes the equity riskier and in turn lowers the demand for the risky asset. The net e ect is that stock return drops more gradually but the total size of crash does not change much.

In contrast, the stock market boom after the disaster is more pronounced in the joint learning case. It is due to the fact that the parameter uncertainty shrinks dramatically through learning during the disaster, so the state uncertainty facing the agents about whether a disaster ends is similar in both cases. In addition, conditional on being in a disaster, the demand for equity is close to that in the partial learning case because of the negligible parameter uncertainty. Conditional on being in the normal time, the demand for equity is lower in presence of parameter uncertainty about future consumption growth so that the equity return is higher. Therefore, in the joint learning case, the equity returns right after the disaster ends are higher than in the case of partial learning.[8]

Putting all the pieces together, the presence of the parameter uncertainty makes equity return in our benchmark joint learning case more volatile than the one in the partial learning case.

[8] At t=13, the return in the joint learning case is lower than the one in the partial learning case. The technical reason for it is because the PDR at t=12 during a disaster with a state belief around 10% is already close to the level of the PDR in normal time. The intuition is that being in a disaster with a not too bad theta and small parameter uncertainty is similar as in the normal state but exposed to much higher parameter uncertainty

Finally, let us turn to compare three cases with joint learning. Because the three cases share the same dynamics of beliefs about both the state and the parameter, the di erences across asset returns should be purely due to how the function for the price-dividend ratio is computed. .

In our benchmark joint learning case, agents are fully rational. They not only acknowledge their imperfect information about but also account for future updates in the belief about when pricing the assets. When the PDR function is computed using the intermediate approach (" x_dist"), agents are myopic in the sense that they ignore changes in their future beliefs and only account for their current belief about the literature on rare disasters.

4.2.1 Moments

Table 2 reports the averages and standard deviations of risk-free rates and levered equity returns. Panel I reports moments from the actual data. We use the U.S. equity return series from the CRSP database, available on WRDS. Data are annual from 1948 to 2009 and expressed in percent. The returns from the actual data are thus corresponding to a consumption process without any realization of disasters.

[Table 2 about here.]

Panel II reports the corresponding model moments. Rows 1 and 2 reports the moments of model-implied asset returns in the case of no learning. Consistent with the sample disaster graphs, parameter uncertainty reduces the risk free rate and raises the equity return. So parameter uncertainty drives up the equity premium but not as much as the one from the actual data. In addition, without any learning, there is little variation in the equity returns compared to the data.

Row 3 shows the results from the partial learning case. Because there is no disaster realization in our simulation and in turn no realization, we assume that agents in this case view equal to the mean of its unconditional distribution F ( ). Learning adds signi cant variation in equity returns compared to the no learning cases. State uncertainty increases equity returns compared to the case with no uncertainty but not as much as what parameter uncertainty does. This case illustrates that a standard Markov model typically requires high risk-aversion and high leverage in order to match equity excess returns.

Row 4 is our benchmark joint learning case. With both parameter and state uncertainty present, the mean of risk free rate is lower and the mean of equity return is raised by a signi cant magnitude, pushing the equity premium closer to the data. The standard deviation of equity returns also improves compared to the partial learning case, and now counts for more than 40% of its counterpart in data.[10] It is worthwhile to emphasize again that those moments are generated

[10]Because the unconditional mean of long-run shock is close to the standard deviation of the shock in normal time, agents are easy to get confused in the partial learning case. In 50% of the time periods, agents think the

without any occurrence of disasters in the simulated sample.

Rows 5 and 6 are the results from two alternative approaches of computing PDR functions in the joint learning case. Borrowing the intuition from the results over a sample disaster, less uncertainty embedded in asset pricing raises the risk free rate and lowers the equity premium. The standard deviation of equity returns increases a lot from our benchmark case due to the large changes in PDRs when switching between low and high consumption growth.

returns.

of dividends can be expected to decline to its true value, leading to lower than expected capital gains along the adjustment path" (p.524)

Some subtle di erences still exist across various learning cases. In particular, R^2 increases when we move do [illegible] over time and the variation in returns comes entirely from the consumption growth.[11] Learning introduces additional volatility in returns, as illustrated by the cases with partial learning and joint learning. In early periods of our sample, the joint learning cases even overshoot the data in terms of return volatility.

Now we take a closer look at the return and belief plots in both the partial learning case and our benchmark joint learning case.

First notice that in both cases, agents belief reaches one during the Great Depression. In the partial learning case, the belief starts relatively low and then peaks up, while in the joint learning case, the belief is already close to one at the beginning of the sample period. Since the return in the initial period is set to be one by construction, the initially peaked belief in the joint learning case limits the change of the price-dividend ratio and that is the reason why the equity return in the joint learning case does not drop as much as in the partial learning case. If data from earlier periods were available, we would see a deeper crash in stock returns in the joint learning case than in the partial learning case.

In the 1930s and 1940s, the belief in the partial learning case reaches 30% during the second World War. However, the dynamics of agents' belief in the joint learning case is mainly determined by two observations of abnormally high consumption growth in mid 1930s (around 10%) and mid 1940s (above 10%). According to the estimation results in Barro et al. (2011), the standard deviation of long-run damage per period during a disaster is as high as 12.1%. Although the parameter of long-run damage, , has a negative mean (-2.4%), the high parameter uncertainty implies a wide range of positive values that can possibly take. Therefore, the large positive consumption growth is viewed by the agents as a rare good event that may last beyond a single p(o)-(erio)-27(de1(v)-339(of)-3133disaster)-439(ou1(v)-31i)1(n)g-439(oarro)-4133dt (2011 and lœk oa

The volatility of consumption growth is significantly lower in the second half of the twentieth century and one can clearly see the great moderation starting in the 80s. Only the recessions in 1970s and early 1990s triggers local spikes in agents' belief in both cases. Nonetheless, we observe larger movements in agents' belief in the partial learning case, consistent with the results in the non-disaster simulation. Moreover, the average belief of being in a disaster is about 5% in the partial learning case, which is more than three times larger than the one in the joint learning case. It may seem counter-intuitive that agents in the partial learning model are easier to confuse a recession with a disaster even though they have perfect information about the disaster parameter. The reason is that with the current setup of consumption process, the long-run damage of a disaster each period is very close to a bad shock in the normal time. This fact helps the partial learning model in fitting the return volatility in the data. A major change in belief occurs at the end of sample in both cases of learning, reflecting the large impact of the 2008 financial crisis. The belief in the joint learning case does not hike up as much as the one in the partial learning case, but it still causes a larger crash in equity returns due to the extra risk brought by the parameter uncertainty.

The bottom two panels in Figure 5 display the model-implied returns in the other two joint learning cases. The return dynamics has similar pattern as the one in the benchmark joint learning case, except missing the two peaks in 1930s and 1940s. It is because the price-dividend ratio during normal time in these two cases are already very high, as compared to the benchmark joint learning case. Thus, the response of price-dividend ratio to the abnormally good event is limited, which in turn mitigates the rise of equity return. Moreover, returns are more volatile over the entire sample due to larger variations of the price-dividend ratio in the state belief, a consistent finding throughout all our results. We can thus conclude that the common approach used in the literature that ignores the parameter uncertainty in asset pricing understates the upward movements but overstates the downward movements in equity returns.

To have a more concrete idea about how the different models fit the data, Table 4 shows the moments of model-implied returns and compares them against the data. Given the return plots, it is not surprising that in the longer sample, our benchmark joint learning model does the best in matching the equity premium and the return volatility of the data. In order to check how much this result is driven by the sample periods in 1930s and 1940s, we show the results of a shorter sample starting from 1948 in Table 5, which excludes the Great Depression and the second World War. Despite the fact that the posterior state belief has much smaller variation than the one in the partial learning case, our benchmark joint learning model still does better in terms of equity premium and return volatility because the parameter uncertainty manages to generate signi cant uctuations in equity returns. This observation underlines the importance of having a uni ed framework to study the state and parameter uncertainty jointly.

[Table 4 about here.]

[Table 5 about here.]

Our results in this section suggests that the performance of a model in matching the return volatility in the data will improve if there are more chances that agents may interpret a bad shock in normal time as the start of a disaster. Intuitively, there should be more confusion when agents are less certain about the parameter. This can be generating by extending the model along the lines of Barro et al (2011), i.e. adding a short-run e ect of a disaster in the consumption process.

4.3.2 Return Pratoctability

Tables 6 and 7 report the results of pratoctive regressions, when the historical data of consumption

In the longer sample, the partial learning model does rather poorly in terms of return predictability in shorter horizons $k = 1; 2$. The slope coe cients are insigni cant and the R^2s are nil. All the cases with joint learning are doing better. The slope coe cient $_k$ are all signi cantly positive and increase with the return horizon. The R^2

In particular, we do not need to rely on the occurrence of disasters or exogenous variations in

References

[1] Abel, A.B., (1990) \Asset Prices under Habit Formation and Catching up with the Joneses." The American Economic Review { Papers and Proceeding, 80(2), 38-42

[2] Adam, K., A. Marcet and J.P. Nicolini, (2008) \Stock MarG.olatility and Learning." European Central Bank Working Paper 862.

[3] Backus, D., M. Chernov, and I. Martin, (2009) \Disasters Implied by Equity Index Options." Working Paper

[4] Bansal, R. and A. Yaron, (2004) \Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." The Journal of Finance, 59(4), 1481-1509

[5] Bansal, R., D. Kiku, and A. Yaron, (2010) \Long Run Risks, the Macroeconomy, and Asset Prices." American Economic Review { Papers and Proceeding, 100(2), 1-5

[6] Barro, R.J., (2006) \Rare Disasters and Asset MarG. in the Tw.ntieth Century." Quarterly

[17] Chen, H. and M. Pakos, (2007) \Rational Overreaction and Underreaction in Fixed Income and Equity Markets: The Role of Time-Varying Timing Premium." MIT Working Paper

[18] Cho, I.K. and K. Kasa, (2011) \Learning and Model Validation." Working Paper

[19] Cogley, T. and T.J. Sargent, (2008a) \The Market Price of Risk and the Equity Premium: A Legacy of the Great Depression?" Journal of Monetary Economics, 55(3), 454-476

[20] Cogley, T. and T.J. Sargent, (2008b) \Anticipated Utility and Rational Expectations as Approximations of Bayesian Decision Making" International Economic Review, 49(1), 185-221

[21]

[35] Jaimovich, N. and S. Rebelo, (2009) \Can News about the Future Drive the Business Cycle?" American Economic Review, 99(4), 1097-1118.

[36] Judd, K.L., (1992) \Projection Methods for Solving Aggregate Growth Models." Journal of Economic Thoery, 58(2), 410-452

[37] Mehra, R. and Prescott, E.C. (1985) \The Equity Premium: A Puzzle" Journal of Monetary Economics, 15(2), 145-161

[38] Piazzesi, M. and M. Schneider (2011) \Trend and Cycle in Bond Premia" Working Paper Stanford University

[39] Rietz, T.A, (1988) \The Equity Risk Premium: A Solution." Journal of Monetary Economics, 22, 117-131

[40] Santa-Clara, P. and S. Yan, (2009) \Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options." Review of Economic Studies, forthcoming

[41]

Table 1: Calibration Parameters

	Symbol	Value
Discount Factor		0.978
Risk Aversion		6.5
IES		2
Av. Consumption Growth		0.022
Mean Disaster		-0.024
Std Disaster		0.121
Std of Cons. Growth Shock		0.018
Leverage		2
Prob. to enter Disaster	p	0.017
Prob. to exit Disaster	1 q	0.153
Learning On	Ton	0.015
Learning O	To	50

Table 2: Asset Pricing Moments

	$E(R_f)$	(R_f)	$E(R^e_{lev})$	(R^e_{lev})	$E(R^e_{lev})$	$E(R_f)$
			Panel I: Data			
	0.97	2.30	8.59	18.03	7.62	
			Panel II: Model			
No learning, no uncertainty	3.04	0.00	4.76	3.76	1.72	
No learning, parameter uncertainty	2.07	0.00	7.08	3.84	5.01	

Table 3: Excess Return Regression:

Lags		1	2	3	4	5
				Panel I: Data		
		0.12***	0.22***	0.27***	0.32***	0.39***
	R^2	0.09	0.16	0.19	0.21	0.25
				Panel II: Model		
No learning, no uncertainty		0.00	0.00	0.01	0.01	0.01
	R^2	0.00	0.00	0.00	0.00	0.00
No learning, parameter uncertainty		-0.01	-0.01	-0.02	-0.02	-0.03
	R^2	0.00	0.00	0.00	0.00	0.00
Partial learning, state uncertainty		0.49***	0.79***	0.97***	1.07***	1.12***
	R^2	0.07	0.10	0.12	0.12	0.12
Joint learning (benchmark)		1.00***	1.15***	1.20***	1.22***	1.22***
	R^2	0.29	0.27	0.23	0.20	0.17
Joint learning, xed belief about		0.89***	1.05***	1.11***	1.12***	1.12***
	R^2	0.36	0.41	0.40	0.39	0.38
Joint learning, xed mean of		0.91***	1.10***	1.17***	1.20***	1.21***
	R^2	0.39	0.46	0.48	0.47	0.46

Notes: The model moments are computed using a long sample with T=100000 periods. The data moments are 1948-2009, * indicates signi cance at the 90% level, ** 95%, *** 99%.

Table 4: Asset Pricing Moments (Historical Consumption Data 1930-2009)

	$E(R_f)$	(R_f)	$E(R^e_{lev})$	(R^e_{lev})	$E(R^e_{lev})$	$E(R_f)$
			Panel I: Data			
	0.49	3.93	7.91	20.03	7.42	
			Panel II: Model			

Table 5: Asset Pricing Moments (Historical Consumption Data 1948-2009)

	$E(R_f)$	(R_f)	$E(R^e_{lev})$	(R^e_{lev})	$E(R^e_{lev})$ $E(R_f)$
Panel I: Data					
	0.97	2.30	8.59	18.03	7.62
Panel II: Model					

Table 6: Excess Return Regression: Historical since 1929

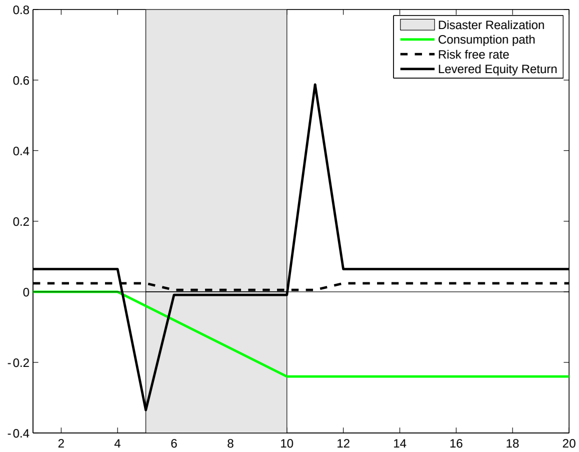
Lags		1	2	3	4	5
				Panel I: Data		
		0.10**	0.21***	0.29***	0.36***	0.43***
	R^2	0.05	0.10	0.16	0.22	0.26
				Panel II: Model		
No learning, no uncertainty		0.00	0.00	-0.01	-0.01	-0.01
	R^2	0.00	0.00	0.00	0.00	0.00
No learning, parameter uncertainty		-0.01	-0.01	-0.02	-0.03	-0.05
	R^2	0.00	0.00	0.00	0.00	0.00
Partial learning, state uncertainty		0.03	0.34	0.98***	1.41***	1.40***
	R^2	0.00	0.02	0.15	0.32	0.33
Joint learning (benchmark)		0.44***	0.75***	1.24***	1.59***	1.59***
	R^2	0.18	0.28	0.52	0.73	0.75
Joint learning, xed belief about		0.27***	0.52***	0.84***	1.15***	1.17***
	R^2	0.17	0.31	0.55	0.83	0.83
Joint learning, xed mean of		0.28***	0.54***	0.87***	1.18***	1.19***
	R^2	0.17	0.31	0.56	0.82	0.83

Notes: * indicates signi cance at the 90% level, ** 95%, *** 99%.

Table 7: Excess Return Regression: Historical since 1948

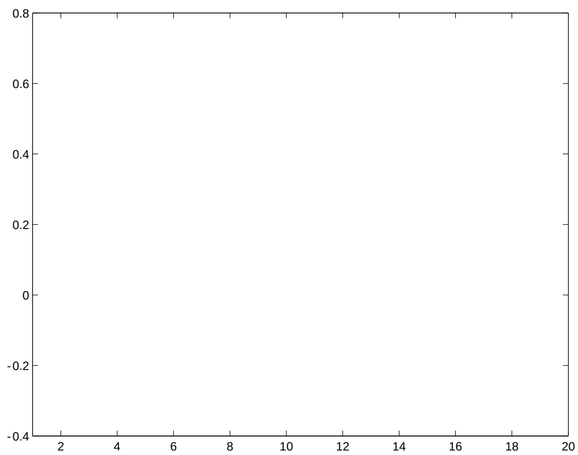
Lags		1	2	3	4	5
				Panel I: Data		
		0.12***	0.22***	0.27***	0.32***	0.39***
	R^2	0.09	0.16	0.19	0.21	0.25
				Panel II: Model		
No learning, no uncertainty		0.00	0.00	0.00	0.00	0.00
	R^2	0.00	0.00	0.00	0.00	0.00
No learning, parameter uncertainty		-0.01	-0.01	-0.01	-0.02	-0.02
	R^2	0.00	0.00	0.00	0.00	0.00
Partial learning, state uncertainty		0.14	0.84*	2.03***	2.37***	2.37***
	R^2	0.01	0.05	0.18	0.22	0.20
Joint learning (benchmark)		0.62***	1.24	2.89***	3.19***	3.11***
	R^2	0.19	0.04	0.14	0.15	0.13
Joint learning, xed belief about		0.44***	0.73	1.66***	1.77***	1.76***
	R^2	0.15	0.04	0.14	0.15	0.15
Joint learning, xed mean of		0.42***	0.86***	1.69***	1.77***	1.72***
	R^2	0.15	0.09	0.21	0.22	0.20

Notes: * indicates signi cance at the 90% level, ** 95%, *** 99%.



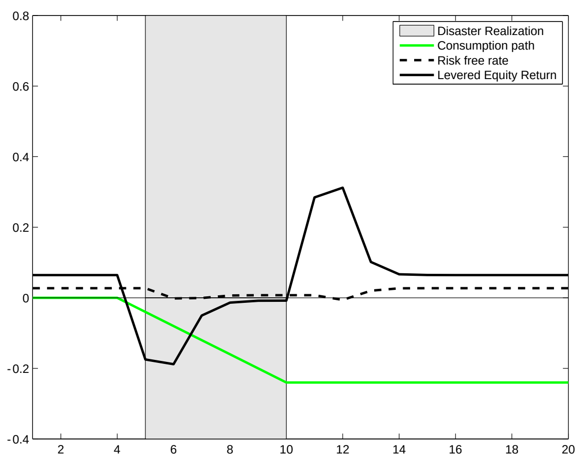
(a) No learning, no uncertainty

(b) No learning, no uncertainty

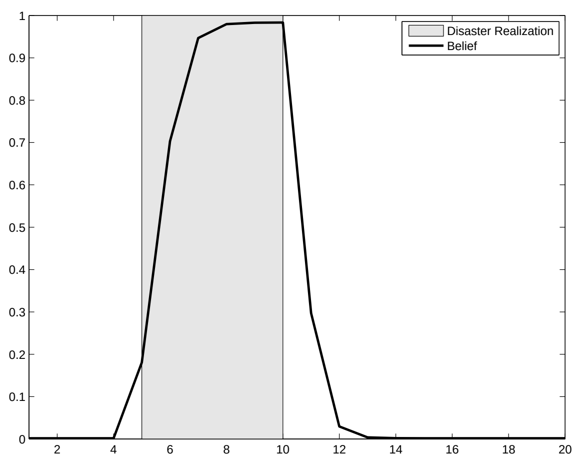


(c) No learning, parameter uncertainty

(d) No learning, parameter uncertainty

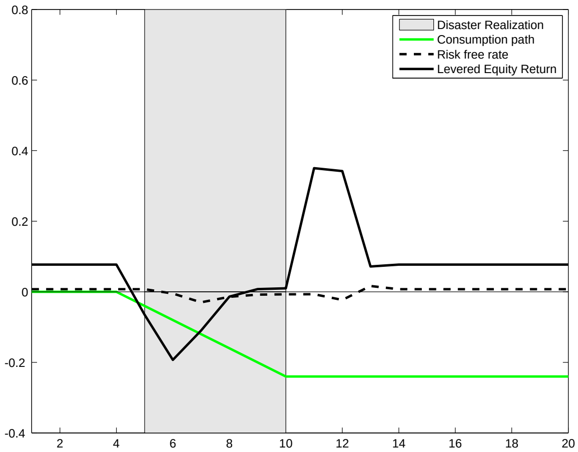


(e) Partial learning, state uncertainty

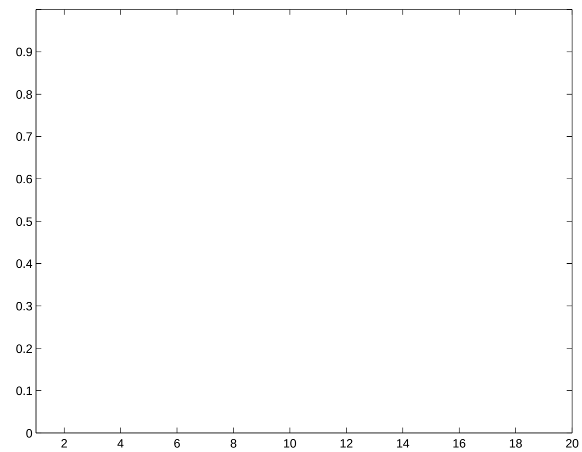


(f) Partial learning, state uncertainty

Figure 1: Sample Disaster, Part 1

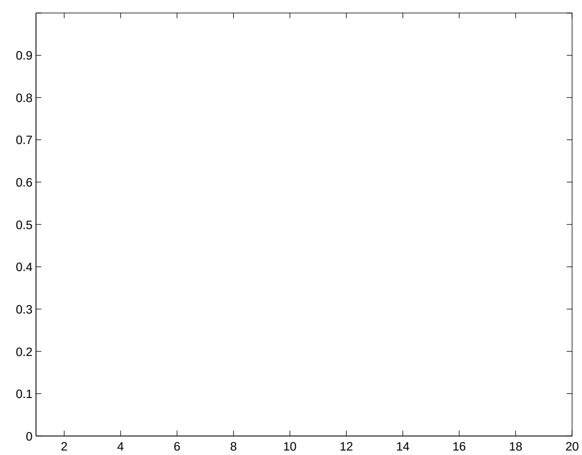


(a) Joint learning (benchmark)

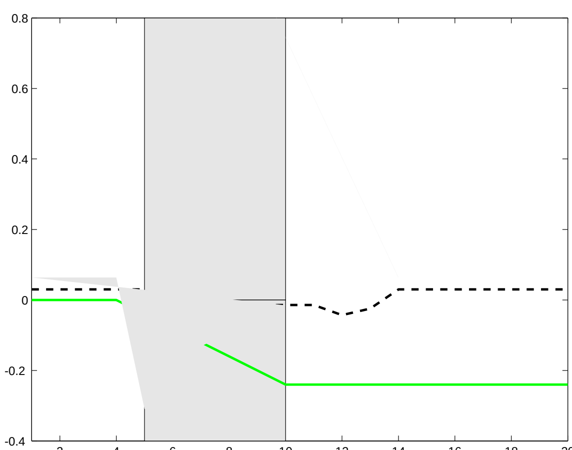


(b) Joint learning (benchmark)

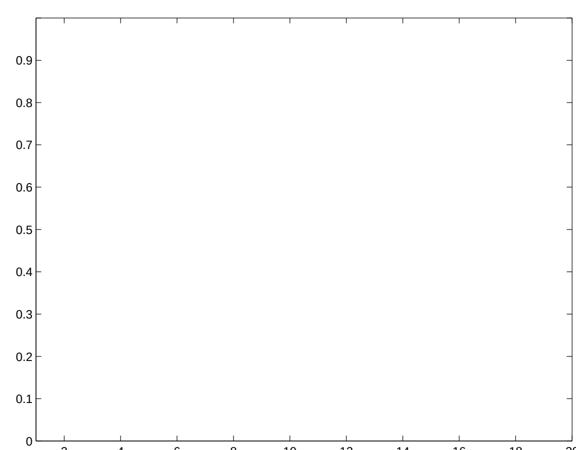
(c) Joint learning, xed belief about



(d) Joint learning, xed belief about



(e) Joint learning, xed mean of



(f) Joint learning, xed mean of

Figure 2:

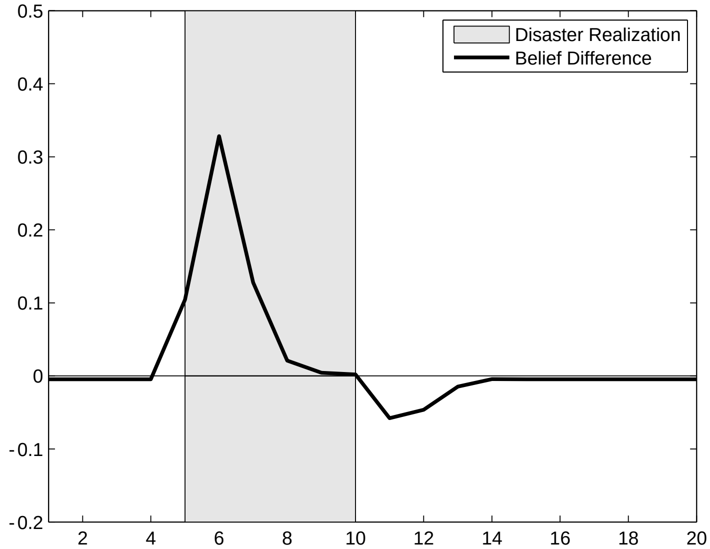


Figure 3: Di erence in state belief: Partial vs. Joint Learning

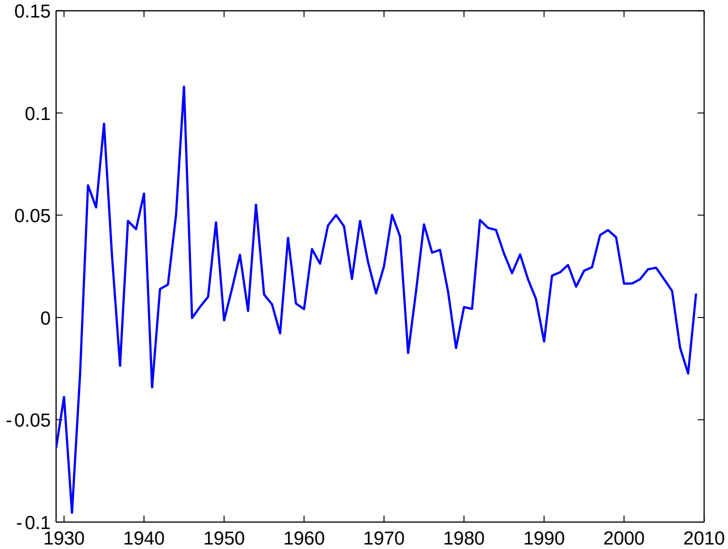
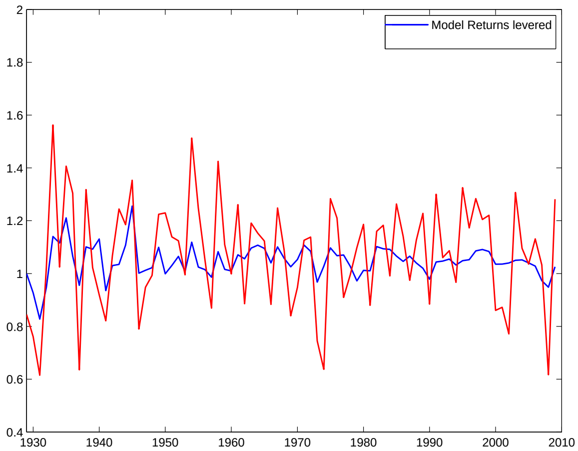
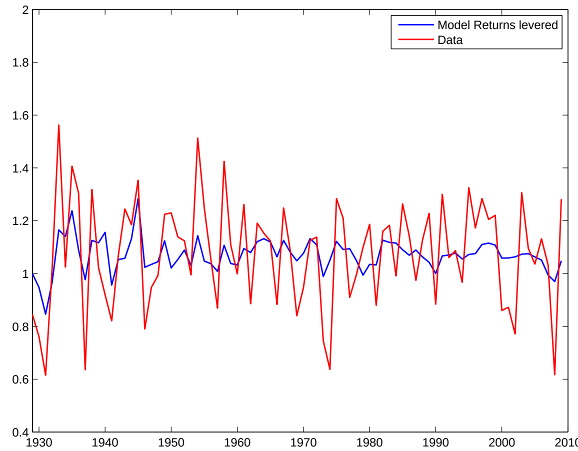


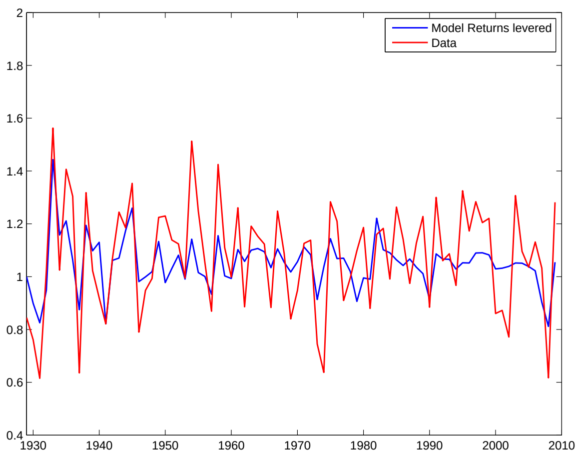
Figure 4: Annual Consumption Growth since 1929



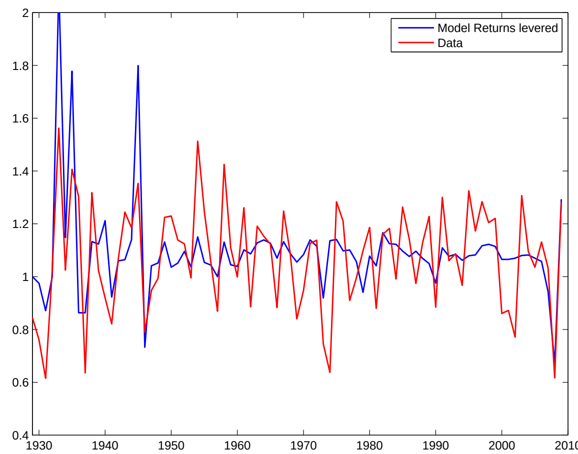
(a) No learning, no uncertainty



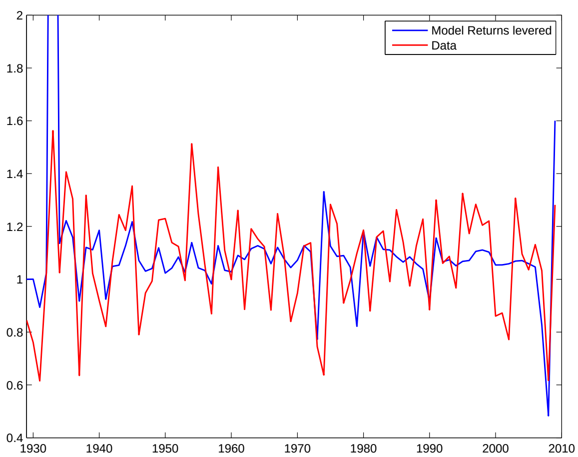
(b) No learning, parameter uncertainty



(c) Partial learning, state uncertainty



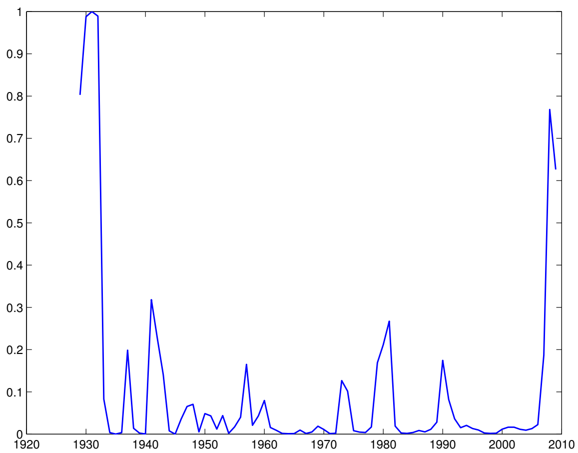
(d) Joint learning (benchmark)



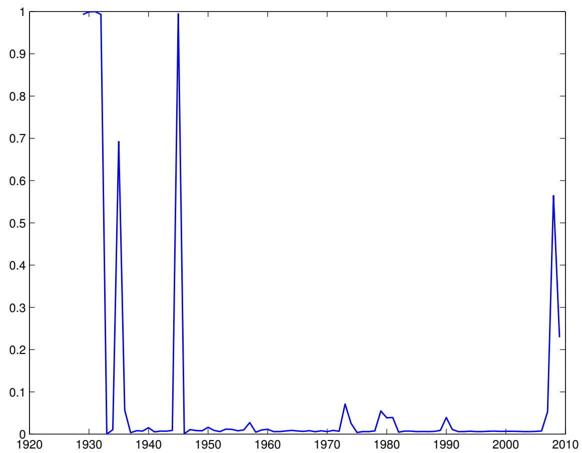
(e) Joint learning, xed belief about

(f) Joint learning, xed mean of

Figure 5: Historical Learning 1929-2009



(a) Partial learning, state uncertainty



(b) Joint learning (benchmark)

Figure 6: Historical Belief 1929-2009