HOOVER INSTITUTION
ECONOMICS WORKING PAPERS

**The Habit Habit**

John H. Cochrane

Economics Working Paper 16105

HOOVER INSTITUTION
434 GALVEZ MALL
STANFORD UNIVERSITY
STANFORD, CA 94305-6010

March 27, 2016

I survey the macro-finance literature related to "By Force of Habit." I show how many models reflect the same rough ideas, each with strengths and weaknesses. I outline how such models may illuminate macroeconomics, by putting time-varying risk aversion, risk-bearing capacity, and precautionary savings at the center of recessions, rather than constraints on flows as in old Keynesian models, or intertemporal substitution and riskfree rate variation as in new Keynesian models. Throughout I emphasize unsolved questions and profitable avenues for research.

# The Habit Habit

John H. Cochrane*

March 27, 2016

**Abstract**

I survey the macro-finance literature related to "By Force of Habit." I show how many models reflect the same rough ideas, each with strengths and weaknesses. I outline how such models may illuminate macroeconomics, by putting time-varying risk aversion, risk-bearing capacity, and precautionary savings at the center of recessions, rather than constraints on flows as in old Keynesian models, or intertemporal substitution and riskfree rate variation as in new Keynesian models. Throughout I emphasize unsolved questions and profitable avenues for research.

## 1. Preface

This talk was prepared for the 2016 "Finance Down Under" conference at the University of Melbourne[1]. I am grateful to the program committee, Carole Comerton-Forde, Vincent Gregoire, Bruce Grundy and Federico Nardari for inviting me, and for selecting my paper with John Campbell, "By Force of Habit" as its "vintage" paper.

As a speech, I recycled some graphs and points from other work, primarily Cochrane (2011) and Cochrane (2007). I also do not pretend to survey the literature evenhandedly, mentioning only a few very specific examples of each approach.

*Hoover Institution, Stanford University and NBER. Webpage: http://faculty.chicagobooth.edu/john.cochrane/research/index.htm.

[1]http://fbe.unimelb.edu.au/conferences/fdu

## 2. Introduction

I am grateful to be invited to reflect on "By Force of Habit," (Campbell and Cochrane 1999a, 199b), though with just a touch of melancholy. How could so many years have passed, with so many projects left undone?

Talks like this about old papers also often turn in to memories of ideas once important, that had their day and have now faded away while we all have gone on to work on other things. I'll try to persuade you that isn't the case; that the research program marked by our habit paper is ongoing and exciting for the future as well as the past. You have seen that before you in the papers in this conference, all cutting-edge research, which in one way or another is pursuing the same agenda.

I was tempted to title my talk something boring, like "habits: past present and future," which is roughly what I'll talk about. But once you start working with habits, they are a bit habit-forming, a habit I will try to pass on. Hence the title.

## 3. A quick habit review

Since much of the audience here was in grade school when John and I wrote "by force of habit," I start with a quick review of the basic idea.

First, we introduce a habit, or subsistence point $X$ into the standard power utility function,

$$u(C) = (C - X)^{1-\gamma}.$$

With this specification, risk aversion becomes

$$-\frac{u''(C)}{Cu'(C)} = \gamma \left( \frac{C}{C - X} \right) = \frac{\gamma}{S}.$$

As $C$ (or the "surplus consumption ratio" $S$) declines, risk aversion rises. (In a multiperiod model, "risk aversion" is properly the curvature of the value function, not the curvature of the utility function. Proper risk aversion turns out to work much the same way in our model.)

Figure 1 illustrates the idea. The same proportional risk to consumption, indicated by the red horizontal arrows, is a much more fearful event when consumption starts closer to habit, on the left in the graph. In the example of the graph, the indicated risk could send consumption below habit, a fate worse than death in this utility function. The same risk, starting at a higher level of consumption, is much more tolerable.

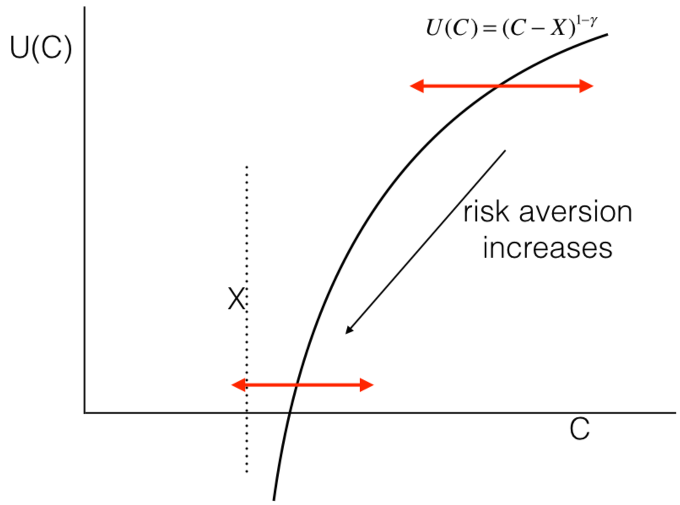


Figure 1: Utility function with habit.

Second, we make the habit slow-moving. Roughly,

$$X_t \approx k \sum_{j=0}^{\infty} \phi^j C_{t-j}; \;\; X_t \approx \phi X_{t-1} + C_t$$

This specification allows us to incorporate growth, which a fixed subsistence level would not do. As consumption rises, you slowly get used to the higher level of consumption. Then, as consumption declines relative to the level you've gotten used to, it hurts more than the same level did back when you were rising. As I once overheard a hedge-fund manager's wife say at a cocktail party, "I'd sooner die than fly commercial again."

Figure 2 graphs the basic idea of the slow-moving habit. As consumption declines toward habit in bad times, risk aversion rises. Therefore, expected excess returns rise. Higher expected returns mean lower prices relative to cashflows, consumption or dividends. Thus a lower price-dividend ratio forecasts a long period of higher returns.

Expected cashflows (consumption growth) are constant in our model, so the variation in the price-dividend ratio is driven entirely by varying risk premiums. Thus, the model accounts for the "excess volatility" of stock prices relative to expected dividends

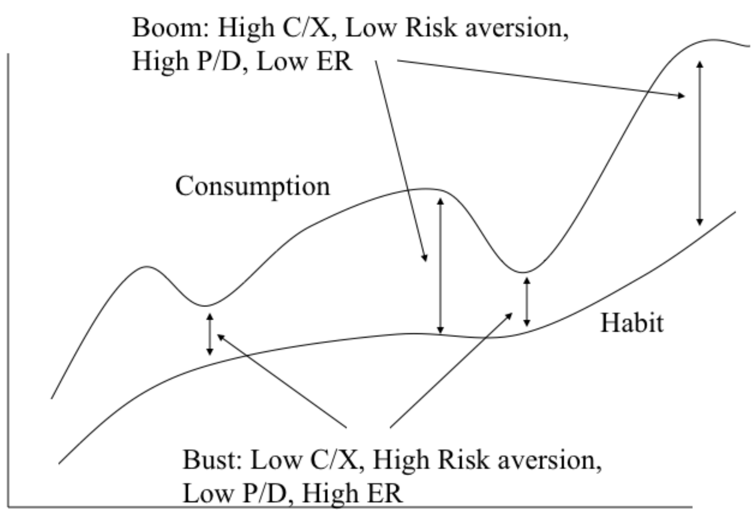


Figure 2: Stylized sample from the habit model.

As Figure 2 illustrates, at the top of an economic boom, prices seem "too high" or to be in a "bubble." But the representative investor in this model knows that expected returns are low going forward. Still, he answers, times are good, he can afford to take some risk, and what else is he going to do with the money? He "reaches for yield," as so many investors are alleged to do in recent years.

Conversely, in bad times, such as the wake of the financial crisis, prices are indeed temporarily depressed. It's a buying opportunity; expected returns are high.

But the average investor looks at this situation and answers "I know it's a good time to buy. But I'm about to lose my job, they're coming to repossess the car and the dog. If it goes down more at all before it rebounds I'm really going to be in trouble. Sorry, I just can't take the risk right now."

In sum, then, as Figure 2 illustrates, the model naturally delivers a time-varying, recession-driven, risk premium. It naturally delivers the fact that returns are forecastable from dividend yields and that forecastability extends to long horizons. It naturally delivers the "excess" volatility of stock prices.

Our model was proudly reverse-engineered. This graph gives our basic intuition going into the project. A note to Ph.D. students in the audience: All good economic models are reverse-engineered! If you pour plausible sounding ingredients in the pot and stir it, you'll never get anywhere.

Third, we engineer the habit accumulation function to deliver a constant interest rate, or in an easy generalization, a slowly varying interest rate.

In a bad time, marginal utility is high, and the consumer expects better (lower marginal utility) times ahead, if not by a rise in consumption, then by a downward adjustment in habit. He would like very much to borrow against that future to cushion the blow today. If he can borrow, that desire would lead to quite persistent consumption growth. If he can't borrow, he will drive up the interest rate in the attempt, and we see strong interest rate variation. The data show neither strongly persistent consumption growth nor large time-variation in real interest rates.

However, in this model, precautionary savings motives are large and time varying. The standard interest rate equation (this is the instantaneous risk free rate with $(C-X)^{-\gamma}$ marginal utility and fixed $X$) is

$$r = \delta + \gamma\left(\frac{X}{C-X}\right) E\left(\frac{dC}{C}\right) - \frac{1}{2}\gamma(\gamma+1)\left(\frac{C}{C-X}\right)^2 \sigma^2.$$

The real interest rate equals the subjective discount factor, plus the elasticity of intertemporal substitution times expected consumption growth, plus risk aversion squared times the variance of consumption growth. As $C-X$ varies, the first term on its own leads either to strong movement in $r$ or in $E(dC/C)$. In standard macro models, risk aversion is low and constant, and variance is much smaller than mean. With $\gamma = 2$ and $\sigma = 0.02$, $\gamma(\gamma+1)\sigma^2 = 2 \times 3 \times 0.02^2 = 0.0024$, a tiny number. But when risk aversion $\gamma/S = \gamma C/(C - X)$ is large, say $25$, to handle the equity premium puzzle, then $25 \times 26 \times 0.02^2 = 0.26$ or $26\%$ on an annual basis. Now precautionary savings matters a lot.

In sum, in bad times, consumers want to borrow against future good times by intertemporal substitution, but they want to save against the possibility of future risk by precautionary savings. Our model exactly offsets these forces to produce a constant risk free rate and iid consumption growth.

That knife-edge is a rhetorical point, not a necessary description of reality. Small changes to the model allow riskfree rate variation and consumption growth variation. We present the knife edge to point out that the model *can* accommodate the extreme case, and thus small generalizations can accommodate reasonable dynamics; that extreme variation of the risk free rate or strong consumption dynamics are not inherent features of a habit model.

All asset pricing models can be expressed as a specification of the stochastic discount factor, the $M$ in

$$1 = E_t(M_{t+1}R_{t+1})$$

or

$$E(R^e_{t+1}) = -cov(R^e_{t+1}, M_{t+1})$$

Assets have higher expected excess returns when they covary more with the discount factor. An asset with a strong negative correlation with $M$ pays off badly when marginal utility is high, when the consumer is hungry, in bad times. The consumer needs a big premium to compensate for that undesirable characteristic.

All of macro-asset pricing comes down to specifying what this $M$ is. What are, exactly, the times or states of nature that investors fear, in which they are hungry, in which cash is particularly valuable; times that the investor would buy insurance to make sure his assets do not fall, foregoing average rate of return to do so?

The standard consumption based model says that consumption growth itself is the purest indicator of such “bad times.” The habit model adds $S$ = the fear of a recession,

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma}.$$

Consumers want to avoid stocks that fall when consumption is low, yes. But with

$\gamma = 2$ this is a small effect. They really want to avoid stocks that fall when $S$ is low – when the economy is in a recession.

As I'll survey in a minute, pretty much all current asset pricing models have this form: they modify the standard consumption based model to include one "extra factor." Consumers are afraid of something else besides consumption. That something else typically also has the flavor of a recession or macroeconomic bad times.

## 3.1. Successes and... room for improvement

So, what does the habit model accomplish?

- Yes: It delivers the equity premium $E(R^e)$ and market Sharpe ratio $E(R^e)/\sigma(R^e)$, with low consumption volatility $\sigma(\Delta c)$, unpredictable $\Delta c_t$, and a low and constant (or slow varying) risk free rate.
- No: It does not have low risk aversion.

Our model really does not "solve the equity premium puzzle." The equity premium puzzle as now distilled includes the equity premium, the market Sharpe ratio, a low and stable risk-free rate, realistic consumption growth volatility, with a positive discount factor $\delta$ and low risk aversion. We have everything but low risk aversion. So far no model has achieved a full "solution" of the equity premium puzzle as stated.

- Yes: The model delivers return predictability, price-dividend ratio volatility, heteroskedastic returns following price declines, and the long-run equity premium.

The model was of course designed to capture long-run return predictability and price-dividend ratio volatility despite iid cashflows. One of its functions has been to point out how those phenomena are really the same.

The long-run equity premium really did pop out unexpectedly after we reverse-engineered much else. Look again at our discount factor

$$M_{t,t+k} = \beta \left( \frac{C_{t+k}}{C_t} \right)^{-\gamma} \left( \frac{S_{t+k}}{S_t} \right)^{-\gamma}$$

Like just about every other model (coming soon), our discount factor adds a second factor that people are afraid of. The equity premium, as distilled by Hansen and Jagannathan (1991), is centrally the need for a higher volatility $\sigma(M_{t,t+k})$ than consumption alone, raised to small powers $\gamma$, provides. The $S$ term provides that extra volatility. In the short run, $S$ and $C$ are perfectly correlated – a positive shock to $C$ raises $C - X$ – so the second factor just amplifies consumption volatility. But in the long run, $S_{t+k}/S_t$ – whether we are in a recession – and $C_{t+k}/C_t$ – the level of consumption – become uncorrelated.

Now, consumption is a random walk, so the standard deviation of the first term rises linearly with horizon. (That's precise only in logs; I'm giving the intuition here of a result that continues to be true in levels.) But our second term, like the second term of all other models, is stationary. The volatility $\sigma(S_{t+k}/S_t)$ eventually stops growing with horizon $k$. If you look far enough out, it would seem, with any stationary "extra factor" you're going to end up with just the consumption model and no extra equity premium.

So, in a robust way, *any* model with a stationary extra factor has a problem, that it does not deliver a rise in $\sigma(M_{t,t+k})$ at long horizons, and thus does not deliver an equity premium at long horizons. Intuitively, temporary price movements really do melt away, so a patient investor collects long run returns and no long-run volatility.

In our model, it turns out that though $S_{t+k}/S_t$ is stationary, $(S_{t+k}/S_t)^{-\gamma}$ is not stationary. Its volatility does increase linearly with horizon, so we have a long-run equity premium puzzle. Marginal utility has a fat tail, a rare event, a min-max or super-salient state of nature that keeps the equity premium high at all horizons. I deliberately use words to connect to the other literature here, as one of my points is the commonality of all the different kinds of models, and the fact that habit models do incorporate many of the intuitions that motivate related models. And vice-versa. However, most of the other explicit models do not capture the long-run equity premium.

How does the model perform since publication? The model says that price-dividend ratios should track the surplus consumption ratio. Figure 3 shows consumption relative to a backward-looking moving average $X_t = k\sum_{j=0}^{\infty}\phi^j C_{t-j}$. (Rather than compute the exact nonlinear model, I chose this more transparent approximation to show that the basic idea is robust.)

As you can see, the brickbats thrown at modern finance for being utterly unable to accommodate the financial crisis are simply false. Consumption relative to habit rises in the pre-crisis boom, and falls at the same time as stock price/dividend ratios fall. The model works better in big events.

Now, for some directions needing improvement. The model has quite a few

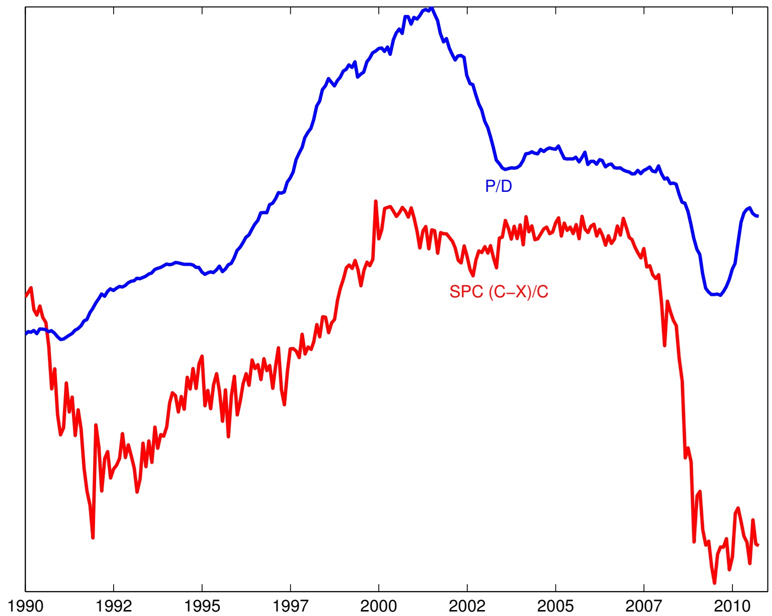


Figure 3: Price-dividend ratio and consumption relative to moving average.

flaws. Most of these flaws are common to alternative frameworks. We expected an active literature that would improve it along these dimensions. That hasn't happened yet, but perhaps I can inspire some of you to try.

- More shocks

The consumption-claim version of our model has one shock, the shock to consumption growth. It is simultaneously a cashflow shock and a discount rate shock, so the cashflow and discount rate shocks are perfectly negatively correlated. When consumption declines (cashflow shock), the discount rate rises.

The standard VAR representation of returns and dividend yields has at least two distinct shocks. In the simplest VAR, cashflow shocks and discount rate shocks are *uncorrelated.*

In round numbers, the standard VAR representation for log returns $r$, log dividend growth $\Delta d$, and log dividend yield $dp$ is

$$\begin{aligned} r_{t+1} &\approx 0.1 \times dp_t + \varepsilon^r_{t+1} \\ \Delta d_{t+1} &\approx 0 \times dp_t + \varepsilon^d_{t+1} \\ dp_{t+1} &\approx 0.94 \times dp_t + \varepsilon^{dp}_{t+1} \end{aligned}$$

and the covariance matrix of the shocks is

$cov(\varepsilon\varepsilon') =$

| | $r$ | $\Delta d$ | $dp$ |
|---|---|---|---|
| $r$ | $\sigma = 20\%$ | +big | -big |
| $\Delta d$ | | $\sigma = 14\%$ | **0 not -1** |
| $dp$ | | | $\sigma = 15\%$ |

The definition of return means that only two of the three equations are needed, and the other one follows. If prices go up or dividends go up, returns must go up! In equations, the Campbell-Shiller return approximation is

$$r_{t+1} \approx dp_t - \rho dp_{t+1} + \Delta d_{t+1}$$

where $\rho \approx 0.96$ is a constant of approximation. As a result of this identity, the VAR regression coefficients $b$ and shocks $\varepsilon$ are linked by identities

$$\begin{aligned} b_r &= 1 - \rho b_{dp} + b_d \\ \varepsilon^r_{t+1} &= -\rho \varepsilon^{dp}_{t+1} + \varepsilon^d_{t+1} \end{aligned}$$

With any two coefficients, shocks, or data series, you can find the last one.

It's common to write the VAR with dividend yields and returns, $\{dp_t, r_t\}$ and let dividend growth be the implied variable. I like to think of it instead in terms of dividend growth and dividend yields $\{dp_t, \Delta d_t\}$ with returns the implied variable. ("Think of it," but don't run it that way. Never ever run a return forecasting regression with less than a pure return.) The reason is that, while $dp$ and $r$ shocks are very negatively correlated – when prices go up, dividend yields go down and returns go up – $dp$ and $\Delta d$ shocks are essentially uncorrelated.

Thus, the easy-to-remember summary of the canonical three-variable VAR is:

- There are two shocks in the data: a cashflow shock $\varepsilon^d$, and a discount rate shock $\varepsilon^{dp}$, and these two shocks are uncorrelated.

The negative correlation of return and dividend yield shocks $\varepsilon^r, \varepsilon^{dp}$, and the positive correlation of return and dividend growth shocks $\varepsilon^d, \varepsilon^r$ then just follows from the last identity.

Clearly, this is a very different picture than our consumption-claim model in which the cashflow and discount rate shocks are perfectly correlated. We need to think of a world with separate and uncorrelated cash-flow and discount-rate shocks, at least when using the dividend yield alone to capture conditioning information.

John and I also had a model with a claim to dividends poorly correlated with consumption, which makes progress towards a two-shock model. Even that model does not replicate the VAR, however. And it suffers from another problem:

- Consumption, stock market value, and dividends are cointegrated.

We just had imperfectly correlated growth rates of consumption and dividends $\Delta c$ and $\Delta d$. But the levels of consumption and dividends wandered away from each other. In the real world, consumption and dividends are both steady shares of GDP in the long run.

Many models have imperfectly correlated $\Delta c$ and $\Delta d$. I have not seen one yet that properly delivers the long run stability of the ratios of stock market value, consumption, and dividends.

- More state variables (?)

Our model has one state variable, the surplus consumption ratio $S_t = (C_t - X_t)/C_t$. The dividend yield is perfectly revealing of this state variable, so no other variable can help to forecast stock returns, bond returns, volatility, or anything else.

The version of our model that allows for time-varying interest rates also has time-varying bond risk premiums forecast by yield spreads. But the bond yield spread is perfectly correlated with the dividend yield so there is effectively only one forecasting (state) variable.

In the model, conditional variances also move around, but again based on the same state variable $S_t$. The conditional Sharpe ratio is not constant, because $E(R^e_{t+1}|S_t)$ and $\sigma(R^e_{t+1}|S_t)$ are different functions.

In the literature, plenty of other variables seem to forecast both stock returns and dividend growth. Martin Lettau and Sydney Ludvigson's (2001) consumption to wealth ratio cay is a good example, which I examined in some depth in "Discount Rates." When we go to the cross-section of returns, size, book-market, momentum, earnings quality and now literally hundreds of other variables are said to forecast returns. Harvey, Liu and Zhu (2016) list 316 variables in the published literature! Bond returns are forecastable by bond forward-spot spreads, and foreign exchange returns by international interest spreads.

Now, a big empirical question remains: Just how many of these state variables do we really need, in a multiple regression sense? The forecasting variables are correlated with each other. Are they both proxies for a single underlying state variable? Or maybe two or three state variables, not hundreds?

The question is, *what is the factor structure of expected returns?* If we run regressions

$$R_{t+1}^i = a_i + b_i x_t + c_i y_t + ..\varepsilon_{t+1}^i;\ E_t(R_{t+1}^i) = a_i + b_i x_t + c_i y_t$$

How many state variables – orthogonal linear combinations of $x, y, z$ – are there? What is the factor structure of $cov\left[E_t(R_{t+1}^i)\right]$? Look at that question closely – this isn't the factor structure of *returns*, time $t+1$ random variables, it's the factor structure of *expected returns*, time $t$ random variables. This covariance and its factor structure may have nothing to do with the factor structure of ex-post returns. Its the factor structure of the linear combinations of forecasting variables that do a good job of forecasting returns, not the factor structure of returns. What is that structure? Across stocks, bonds, foreign exchange etc.? As a hint, Monika Piazzesi and I (2005, 2008) found that the covariance of bond expected returns across maturities has one very dominant factor. Does that observation extend to bonds and stocks? Probably not. But our bond-forecasting factor forecasts stocks, and dividend yields forecast bonds. How much of a second factor do we really need? Bringing some order to the zoo of factors that forecast the cross-section of stock returns is even more important – I hope we don't need 300 separate factors.

Conditional variances $\sigma_t(R_{t+1})$ vary over time as well. The empirical literature seems to focus on realized volatility – lagged squared returns – and volatilities implied by options prices as the state variable for variance. These variables decay much more quickly than typical expected return forecasters like dividend yield. Realized volatility also forecasts mean returns, though, and dividend yields forecast volatility. How many state variables are there really driving means and variances?

Finding the factor structure of conditional moments (mean and variance), and seeing how many different forecasters we really need, is a big and largely unexplored empirical project.

The answer is unlikely to be one, as specified in our model. Hence, the natural generalization of theory must be to include more state variables, to match the more state variables in the data. Jessica Wachter (2006) has taken a step in this direction, separating somewhat bond and stock forecasts, but there is a long way to go.

Finally, there is a flurry of work now looking at the term structure of risk premiums, which may provide a new set of facts for models to digest. In simplest form, this work distinguishes $E_t R_{t+k}$ across different horizons $k$. In my evaluation the empirical facts of this literature are still tenuous for solid model fitting, but the direction of research is worth noting.

- Tests

Habit models really have not been subject to much formal testing. (Tallarini and Zhang 2005 is a lonely counter example.)

Of course, as we are learning with the second generation of consumption based model tests, those glasses can be a lot more full than we thought. Many of the early rejections used monthly, seasonally adjusted, time-aggregated consumption data. No surprise that didn't work. More recent tests, such as Jagannathan and Wang's (2007) use of fourth quarter to fourth quarter annual data, find unexpected success for the consumption based model. Our theoretical model also showed how time-aggregation could destroy model predictions.

So we're all waiting, really, for a really good assessment of the consumption based model with habit and other novel preferences, but doing its best to see where the glass is half full, by treating durability (nondurable consumption includes clothes for example), seasonality, time aggregation, and so forth.

Our full model can swiftly be rejected. All explicit economic models have $R^2 = 1$ predictions in them somewhere, unless the researcher salts them up with measurement error shocks. The permanent income model says consumption is the present value of future income, with no error term. The Q theory of investment says that investment = a function of stock prices, with no error term. Our model says that the dividend yield is nonstochastic function of the surplus consumption ratio. A graph such as Figure 3 is a 100% probability rejection of the model, because the consumption and stock price lines are not exactly on top of each other to the 18th decimal point. So the real art of testing is to see in what sensible predictions of a model are really at odds with the data, avoiding "rejecting" a model because a 100% $R^2$ prediction is only 99.9% in the data.

- Low hanging fruit for *all* similar models.

These deficiencies are common to all of the class of macro-asset pricing models. I list them as partial defense against the old-paper all-played-out syndrome. There is lots of low-hanging fruit in this business!

## 4. Other directions

The literature did not follow this roadmap. To be honest, the following years have not seen a flowering of research using the habit model.

Instead, the finance or macro-finance literature explored alternative preferences and market structures to much the same ends as we did. A small sampling:

1. Recursive utility (Epstein and Zin 1989).
2. Long run risks (e.g. Bansal Yaron 2004; Bansal Kiku Yaron 2012).
3. Idiosyncratic risk (e.g. Constantinides and Duffie 1995).
4. Heterogeneous preferences (e.g. Garleanu and Panageas 2015).
5. Rare Disasters (e.g. Reitz 1988; Barro 2006).
6. Nonseparable across goods (e.g. Piazzesi, Schneider and Tuzel 2007).
7. Leverage; balance-sheet; "institutional finance" (e.g. Brunnermerier 2009, Krishnamurthy and He 2013, many others).
8. Ambiguity aversion, min-max preferences, (e.g. Hansen and Sargent 2001 and following).
9. Behavioral finance; probability mistakes (e.g. Shiller 1981, 2014.).

And many more.

These approaches look different, but in the end the ideas are quite similar. Each of them boils down to a generalization of marginal utility or discount factor looking much like ours,

$$M_{t+1} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} Y_{t+1}$$

The new variable $Y_{t+1}$ does most of the work, and varies over time with recessions. Even the behavioral and probability distortion views are basically of this form. Expressing the expectation as a sum over states, the basic first order condition is

$$P_t U'(C) = \beta \sum_s \pi_s(Y) U'(C_s) X_s$$

Probability and marginal utility always enter together, so distorting marginal utility by adding another variable is the same thing as distorting probabilities. The state variables $Y$ driving probability distortions act then just like state variables driving marginal utility.

The ideas are in fact quite similar, as I'll stress in a lot of contexts. Thus, rather than a grumpy "you all should go back and use habits," my real point is that one can use almost any of these formalisms to capture roughly the same ideas.

Which formalism we use in part depends on which data for $Y$ turn out to work best. But most of the candidates are highly correlated with each other, so telling them apart will be hard and thus less important than it seems. Which one we use may end up therefore depending even more on simple tractability. I'll make the case that the habit formulation is at least as convenient as the others, has already captured more phenomena more elegantly – with fewer auxiliary assumptions– as the others, and therefore is at least worth playing in the same field. But that's not a huge advantage, and in fairness a more tractable modeling formalism may appear that improves over all of these.

To make that point, I examine a few of the alternatives in some detail.

## 4.1. Recursive utility / long-run risk

The recursive utility approach uses a nonlinear aggregator to unite present utility and the future,

$$U_t = \left( (1-\beta) c_t^{1-\rho} + \beta \left[ E_t \left( U_{t+1}^{1-\gamma} \right) \right]^{\frac{1-\rho}{1-\gamma}} \right)^{\frac{1}{1-\rho}} .$$

Here $\gamma$ is the risk aversion coefficient, and $1/\rho$ is the elasticity of intertemporal substitution. It reduces to power utility for $\rho = \gamma$.

The discount factor, or growth in marginal utility, is

$$
\begin{aligned}
M_{t+1} &= \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\rho} \left( \frac{U_{t+1}}{\left[ E_t \left( U_{t+1}^{1-\gamma} \right) \right]^{\frac{1}{1-\gamma}}} \right)^{\rho-\gamma} \\
&= \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\rho} \left( Y_{t+1} \right)^{\rho-\gamma}.
\end{aligned}
$$

In the latter equation, I emphasize that the innovation in the utility index takes the role of the new variable $Y$ in my general classification.

The utility index itself is hard to observe, so the trick is to substitute for it in terms of other, more observable variables. Lately, the most common approach, exemplified by Bansal, Kiku, and Yaron (2012), and Hansen, Heaton and Li (2008), has been to substitute the utility index for the stream of consumptions that generate utility, which delivers the long run risk model. For $\rho \approx 1$,

$$
\Delta E_{t+1} \left( \ln M_{t+1} \right) \approx -\gamma \Delta E_{t+1} \left( \Delta c_{t+1} \right) + (1-\gamma) \left[ \sum_{j=1}^{\infty} \beta^j \Delta E_{t+1} \left( \Delta c_{t+1+j} \right) \right]
$$

where $\Delta E_{t+1} \equiv E_{t+1} - E_t$.

Now the news about long run future consumption growth is the extra state variable. As usual this extra state variable will have to do the bulk of the work to explain risk premiums. So, this is a model like habits, in which people are afraid of a state variable in addition to consumption growth, and people are afraid of stocks that covary with this new state variable. The innovation: people are afraid of stocks that might go down *when there is bad news about long-run future consumption growth*, not necessarily in a recession, or a time when consumption is low relative to its recent past.

I note with some chagrin that the recursive utility/long run risk framework is much more popular than our habit persistence. Still, I think there is room to question the wisdom of this popularity.

First, the model crucially needs there to *be* news about long run consumption growth – variation in $\Delta E_{t+1} \Delta c_{t+j}$, $j > 1$ – to get anywhere. If consumption is uncorrelated over time, if day by day consumers answer a hypothetical survey about their expectations of consumption growth (not level) in 2030 with the same number, say 1%, then there is no long-run consumption news and the model reduces to power utility.

Current conditions $\Delta c_t$ are essentially irrelevant to investor's fear. Investors only seem to fear stocks that go down when current consumption goes down (fall 2008, say) because, by coincidence, current consumption declines are correlated with the bad news about long run future consumption that investors really care about.

So is there a lot of news about long-run consumption growth? And is it at all believable that this is really what investors care about? The former is hard to find in the data. Apart from a first-order autocorrelation due to the Working effect (a time-averaged random walk follows an MA(1) with an 0.25 coefficient) and the effects of seasonal adjustment (our data is passed through a 7 year, two-sided bandpass filter), nondurable and services consumption looks awfully close to a random walk. (Beeler and Campbell's 2012 review is good on this point.) Inferring long-run predictability from a few short-run correlations is a dubious business.

One might retort, well, the standard errors are big, so you can't prove there *isn't* a lot of very long run positive autocorrelation in consumption growth. (My own "Random Walk in GNP," Cochrane (1988), written even longer ago, was exactly about the difficulties of measuring lots of small high order autocorrelations.) But demoting the central ingredient of the model from a robust feature of the data to an assumption that is hard to falsify clearly weakens the whole business.

I often advise students to write the op-ed or teaching note version of their paper. If you can't explain the central idea to a lay audience in 900 words, or if it is embarrassingly thin when you try, then maybe it isn't such a good idea after all.

In this case, that oped would go something like this: Why were people so unhappy with their stocks in, say, winter 2008? It was not, really, because the economy was in a recession, that investors had lost their jobs and houses and they were cutting back on consumption. That fact, per se, was irrelevant. Instead, it was because 2008 came with bad news about the long-run future. Investors figured out what no professional forecaster did, that we would enter these decades of low growth. If that bad news about long run growth happened to be correlated with a boom in 2008, stocks would nonetheless have collapsed and people would have paid dearly ex ante to avoid stocks that did particularly badly on that news. People, and the institutions such as university endowments trying to sell in a panic, didn't fundamentally care *at all* about what was happening in 2008 – it's only the long run news that mattered to them. This strikes me as a difficult essay to write, or a difficult proposition to explain honestly to an MBA class on any day but the first of April.

To understand the long-run risk model, ask this (a good exam question): How is the long-run risk model different from Merton's ICAPM? After all, the ICAPM also includes additional pricing factors, that are "state variables for investment opportunities." News about long-run consumption growth would certainly qualify as an ICAPM state variable. Yet the ICAPM has power utility. Why did we need recursive utility? (Pause. Long silence.)

The answer is that the ICAPM is a subset of the power-utility consumption-based model. Its multiple factors are the market return and state variables, not consumption and state variables. If one observes consumption, the ICAPM reduces to the single-factor consumption model. In response to bad news about future consumption, ICAPM consumers reduce consumption today. That reduction in today's consumption reveals all we need to know about how much the bad news hurts. The bad news matters, but does not enter as a separate state variable in addition to current consumption.

By contrast, the long-run risks model weights news about future consumption that is *not* reflected in consumption today. Somehow, you get news that the future is a disaster, yet you may still choose to live high today. *This* is the kind of bad news about which you are really afraid. If you did react by lowering consumption today, then consumption would be a random walk and we'd be back to power utility.

As you can tell, this all sounds pretty thin to me. In the habit model, people really are worried about stocks falling in 2008 – because of events going on in 2008! They are worried because consumption falls when it is already near habit, defined as a fraction of the level of consumption they have gotten used to in the previous decade.

This behavior is closely related to the central theoretical advertisement for recursive utility. Recursive utility captures – and requires – a "preference for early resolution of uncertainty." This is a tricky concept. In almost all of your experience you prefer to resolve uncertainty early because you can do something about it. If you know what your salary will be next year, you can start looking for a better house, or a better job. If you learn what the stock market will do next year, you can buy today. The preference for early resolution of uncertainty that these preferences capture is a preference for learning the future, even when you *can't do anything* in response to the news.

I find lab experiments documenting such preference unpersuasive, because there is essentially no circumstance in daily life in which one gets news that one can do absolutely nothing about. People respond to surveys and experiments with rules of thumb adapted to the circumstances of their lives.

Larry Epstein, Emmanuel Farhi, and Tomasz Strzalecki (2014) address the question this way: How much would the consumer in the Bansal-Yaron economy pay, by accepting a lower overall level of consumption, in order to know in advance what consumption will be? The answer is around 20 to 30 percent – the consumer would accept a stream that is 20 to 30 percent lower on every single day of his or her life, just for the psychic pleasure of knowing what it will be in advance. That seems like a lot.

One real-world circumstance that almost fits the model is genetic testing for Huntington's disease. There is no cure, you simply find out if you're going to get the disease. In this case there is quite a bit one can do with the information, such as make career, family, investment, and estate decisions. Nonetheless, , however, Emily Oster, Ira Shoulson and E. Ray Dorseyvery (2013) point out that few people get the test.

So capturing a strong preference for early resolution of uncertainty starts to me to look more like a bug than a feature.

This isn't some mongrel unrelated issue – it's central to the whole long-run risks idea. The news about future consumption, unrelated to current consumption, that so drives risk premiums in the model, is exactly this psychic pleasure or pain of learning the future, unrelated to current action or any planning, investing, or other actions one can take in regard to the news. If you don't believe one, you don't believe the other.

The other apparent theoretical advantage is that recursive utility "separates risk aversion from intertemporal substitution," allowing high risk aversion for the equity premium and a low and steady risk free rate.

But so do habits! Here I grant that recursive utility achieves the result more elegantly. In the habit model, we delicately offset time-varying intertemporal substitution demands with a time-varying precautionary saving. Elegance and tractability are really important in economic theories. Unlike commentators who decry too much math in economics, our problem is too little math – we don't have enough simple tractable mathematical models to play with and see how things hang together. The lesson of the success of recursive preferences is important for model-builders: tractability is more important than realism. As the old joke goes, we look for our car keys under the street lamp, where it's light, not down the dark street where we dropped them.

But that elegance and tractability may lead us astray. If in fact time-varying precautionary saving is important – if, say, Fall 2008 had a large fall in consumption because people were scared to death – then the model is missing the crucial feature of reality. Furthermore, though people complain that the square root habit adjustment process in our model is complex, really, the complexity is nothing like the algebra one must go through to solve recursive utility models.

The recursive utility model, like the habit model, produces the equity premium with a low and stable risk free rate and realistic (low) one-period consumption volatility. It can use high risk aversion, as we do. It can also produce the equity premium with relatively low risk aversion, by imagining a lot of positive serial correlation in consumption growth – a lot of long-run news. In this case, though, long run consumption volatility is very high, so it is in the class of theories that abandon the low consumption volatility ingredient of the equity premium puzzle statement.

But the more interesting, and challenging, phenomena are return predictability and time-varying volatility. The long-run risk model does not naturally produce time-varying risk premia. These have to be put in by assuming an exogenous pattern of consumption volatility. This explanation of predictability goes back to Kandel and Stambaugh in the late 1980s with power utility: To get $E_t(R^e)/\sigma_t(R^e) \approx \gamma\sigma_t(\Delta c_{t+1})$ to vary over time, you need to imagine that $\sigma_t(\Delta c_{t+1})$ varies over time.

Again there is very little direct evidence for that proposition. Moreover, it's another exogenous coincidence. The habit model builds in a time-varying Sharpe ratio, higher in bad times, *endogenously*. Risk aversion $\gamma$ rises as consumption falls towards habit.

So, all the interesting predictions of the model have to be baked in by the assumptions on the exogenous consumption process. As a result, the predictions are very sensitive to those assumptions, and there is little clear direct support for those assumptions in data. And it raises the question whether in a production and investment economy, consumers will choose a consumption process with just the right correlation of short-run and long-run risks, and the variation in volatility, needed to produce the large asset pricing swings we observe.

Hui Chen, Winston Dou and Leonid Kogan (2015) charge that many asset pricing models are based on what they humorously call "dark matter." Unobserved state variables drive marginal utility, the discount factor, or probability assessments. Absent some independent measurement, "Markets went up because long-run news was good," "markets went up because risk aversion declined," "markets went up because sentiment rose," "markets went up because the chance of a rare disaster declined" are no better really than, "markets went down because the Gods are angry." Ok, that sort of analysis can "explain" any phenomenon. But we'd rather the number of assumptions were less than the number of predictions.

To progress somewhere, any extra state-variable model needs to propose some *independent* way of measuring shifts in marginal utility, and that measurement should contain as few extra assumptions as possible. In the habit model, the extra state variable – surplus consumption ratio – is directly and independently (from asset prices) measurable. Furthermore, it generates the extra state variable – surplus consumption ratio – *endogenously* via the link between consumption and habit.

The Bansal-Yaron model ties its dark matter – news about long run consumption growth – to observable data by the assumption that short-run consumption growth and its volatility are correlated with the long-run news. That assumption makes long-run news independently measurable. But the crucial link is driven by extra assumptions about the exogenous driving process, having nothing to do with the preferences.

However, I want to emphasize the inclusive note. Both models capture a quite similar idea. There is an extra state variable, which explains why people are afraid of holding stocks in ways not described by just consumption growth. That extra state variable has something to do with recessions, bad macroeconomic times. Both models capture an equity premium and time-varying predictability, one with time-varying risk, the other with time-varying risk aversion. There is not a huge difference. No model has gotten significantly ahead of the others in terms of the number of phenomena it captures. All models have inconvenient truths that we ignore, as the original CAPM required no investor to hold a job, and predicted that consumption volatility is the same as market volatility. That didn't stop it from being a useful model for many years.

For reference: The Bansal Yaron Kiku consumption process is

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \sigma_t \eta_{t+1} \\
x_{t+1} &= \rho x_t + \phi_e \sigma_t e_{t+1} \\
\sigma_{t+1}^2 &= \bar{\sigma}^2 + v(\sigma_t^2 - \bar{\sigma}^2) + \sigma_w w_{t+1} \\
\Delta d_{t+1} &= \mu_d + \phi x_t + \pi \sigma_t \eta_{t+1} + \phi \sigma_t u_{d,t+1}
\end{aligned}$$

## 4.2. Idiosyncratic risk

Idiosyncratic risk, made famous in asset pricing by Constantinides and Duffie (1996), is another fundamentally different microeconomic story that also generates similar results.

The bottom line is again a discount factor that adds a state variable to consumption growth,

$$M_{t+1} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( e^{\frac{\gamma(\gamma+1)}{2} y_{t+1}^2} \right).$$

Here $y_{t+1}$ denotes the *cross-sectional variance* of consumption growth. The log of each individual's consumption follows

$$\Delta c_{t+1}^i = \Delta c_{t+1} + \eta_{i,t+1} y_{t+1} - \frac{1}{2} y_{t+1}^2; \; \sigma^2 \left( \eta_{i,t+1} \right) = 1$$

Therefore, $y_{t+1}$ plays the role of the second, recession-related state variable in place of the surplus consumption ratio or long-run risk.

The story: People are afraid that stocks might go down at a time when they face a lot of idiosyncratic risk. Some might get great gains, some might face great losses. With risk aversion, i.e. nonlinear marginal utility, fear of the losses outweighs pleasure at the gains, so overall people fear assets that do badly at times of great idiosyncratic risk.

The Constantinides and Duffie paper is brilliant because it is so simple, and it provides directions by which you can reverse-engineer any asset pricing results you want. Just assume the desired cross-sectional variance $y_{t+1}$ process. (It also circumvents many problems with the previous literature.)

As with the long-run risks model, however, the level and especially the time-variation and business cycle correlation of the equity premium all are baked in by the *exogenous* variation in the moments of the income process, rather than the endogenous response of risk aversion to bad times. Cross-sectional consumption volatility must be large, and must vary a good deal over time, and at just the right times.

One can check the facts, and so far the empirical work has been a bit disappointing to the model. Cross-sectional risks do rise in recessions, and when asset prices are low, but they do not seem large enough, or time-varying enough to generate the asset pricing phenomena we see at least with low levels of risk aversion. Consumption risks are much smaller than transitory income or employment risks. However, this is still an active area of empirical research. For example, Lawrence Schmidt (2015) has recently investigated whether the non-normality of idiosyncratic risks can help – whether a time-varying probability of an idiosyncratic rare disaster dominates the cross-sectional risks to marginal utility. Such events are intuitively plausible.

Again you can see the essential unity of the ideas. A second state variable, associated with recessions, drives marginal utility. People are afraid that stocks might fall in recessions, and being in a recession and a time of low-price dividend ratios raises that fear. Here "recessions" are measured by a large increase in idiosyncratic risk, rather than by a fall of average consumption relative to its recent past. But those events are highly correlated. I admire the elegance that the second state variable is endogenous, and tied directly to the fall in consumption, rather than exogenous and requiring an extra set of assumptions. But that too is minor. The moments of cross-sectional risk are at least more tightly tied to data than the inference about long-run risk from its correlation with short run risks.

## 4.3. Heterogeneous preferences

Garleânu and Panageas (2015) offer a related but diametrically opposed model. For Constantinides and Duffie, people have the same preferences, risks are not insured across people, and exposure to this time-varying cross sectional risk drives asset prices. For Garleanu and Panageas, people have different preferences. Some are more risk averse, and some are less risk averse. The risks are perfectly insured across people. Now, less risk averse people hold more stock than more risk averse people. But, when the market goes down, the big stockholders lose more money, and so they become a smaller part of the overall market. So, by shifting consumption from the risk-takers to the risk-haters, the market as a whole becomes more risk averse after a fall in value.

More precisely, in a complete market the unique discount factor $\Lambda_t$ and consumer $A$, $B$, consumption follow

$$\Lambda_t = e^{-\delta t} c_{A,t}^{-\gamma_a} = e^{-\delta t} c_{B,t}^{-\gamma_b}$$

Thus in bad times, with high $\Lambda_t$, the less risk averse consumer accepts greater consumption losses, while in good times, that consumer enjoys greater gains. Mechanically, this sensitivity is implemented via greater investment in the market.

Differentiating these relationships, we can express the discount factor in terms of aggregate consumption $c_t = c_{A,t} + c_{B,t}$ raised to an aggregate risk aversion, which is the consumption - weighted average of individual's inverse risk aversion.

$$\frac{1}{\gamma_{mt}} = \frac{1}{\gamma_B} \frac{c_{B,t}}{c_t} + \frac{1}{\gamma_A} \frac{c_{At}}{c_t}.$$

You see here exactly the sort of mechanism of a habit model – the representative agent becomes more risk averse after a fall in value. But here, it is not because each *individual* becomes more risk averse, as it is in our habit model. It is because the mechanism of aggregation puts more weight on the risk averse people in bad times.

This is a beautiful model, which emphasizes just how many micro stories are consistent with the same macro phenomenon. Perhaps we should not say that markets become "more risk averse," but that they display "less risk bearing capacity" in bad times. That phenomenon can be driven by market structures as well as by psychology of individual preferences. This model faces challenges in the micro data just as the idiosyncratic risk model does. Do the "high-beta rich" really lose so much in bad times? But that investigation hasn't really started.

### 4.4. Balance sheets, debt, and institutional / intermediated finance

A different category of model has become much more popular since the 2008 financial crisis: models involving debt, balance sheets, mortgage overhang, and "institutional finance."

The basic story works much like habit persistence. Imagine that an investor has taken on a level of debt $X$, which he must repay. Now, as income declines towards $X$, he will take on less and less risk, to make sure that even in bad states of the world he can repay his debt. The intuition of Figure 1 applies exactly, if we just re-label $X$ as the level of debt.

Moreover, as consumption rises in good times, people slowly take on more debt. As consumption falls in bad times, people "delever," "repair balance sheets" and so forth. So debt moves slowly, following consumption, very much like our slow-moving habit.

It's not so easy, however.

First, why do agents get *more* risk averse as they approach bankruptcy, not *less*? Bankruptcy is the point at which you don't have to pay your debts any more. It is usually modeled as a call option. The usual concern is that people and businesses near bankruptcy have incentives to take *too much* risk, not too little. If the bet wins, you're out of trouble. If the bet loses, the bank or creditors take bigger losses – not your problem.

The costs, benefits, reputational concerns, and so forth surrounding bankruptcy are subtle, of course, and I don't mean to argue that we know exactly one way or another in all circumstances. I do point out that it's not at all obvious that debt should induce more risk aversion rather than less, and it takes modeling effort and dubious assumptions to produce the more answer.

Second, not everyone is in debt. My debt is your asset – net debt is zero. For this reason, institutional finance models center on segmented markets, so that the problems of borrowers weigh more heavily on markets than the problems of their creditors.

The typical institutional finance story told of the financial crisis goes like this, illustrated in Figure 4. Fundamental investors – you and me – give our money to intermediaries. The intermediaries take on leverage, so in essence we split our funding of the intermediaries into debt and equity tranches. When the intermediaries start losing money, they get more risk averse, and start selling assets. (For various reasons they don't raise more equity, give us securities, or bet the farm on riskier trades.) You and I don't trade actively in the underlying assets so there is nobody around to sell to. Only the intermediaries are "marginal." Hence, when they try to sell, prices go down. That puts them closer to bankruptcy, so they sell more, with colorful names like "liquidity spiral," or "fire sale."

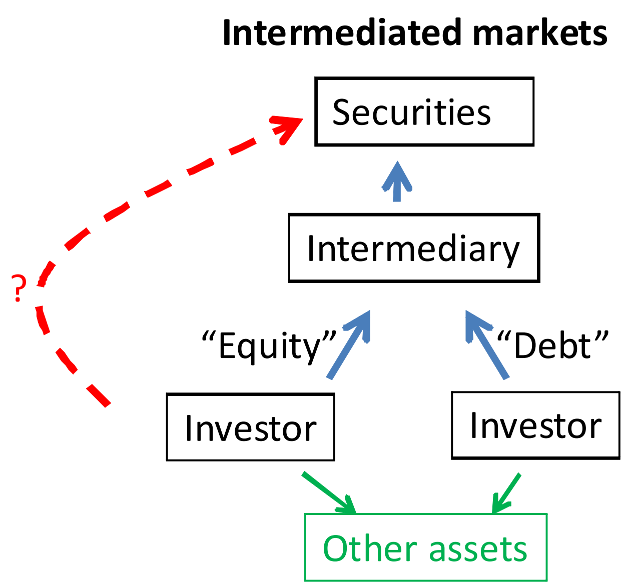


Figure 4: Schematic of intermediated asset pricing.

The objections to this sort of model are straightforward. OK, for obscure CDS or other hard to trade instruments, and this may explain why small arbitrages opened up between more obscure derivatives and more commonly traded fundamentals. But how does this story explain the widespread, coordinated, falls in stock and bond markets around the world? After all, these assets are part of everybody's pension funds. We're all "marginal." Moreover, large, sophisticated, unconstrained, debt-free wealthy investors and institutions such as university endowments, family offices, sovereign wealth funds, and pension funds all trade stock indices and corporate bonds every day. If leveraged intermediaries push prices down nothing stops these investors from buying. (I suggest this direct linkage with the dashed red line.) Where were they?

Answer: they were selling in a panic like everyone else. That surely smacks of time-varying risk aversion, induced by recent losses, not a segmented market in which fundamental investors want to buy but leverage and agency problems cause their agents to sell.

Furthermore, if there is such an extreme agency problem, that delegated managers were selling during the buying opportunity of a generation, why do fundamental investors put up with it? Why not invest directly, or find a better contract?

To emphasize the coordination of asset price falls in many different markets, Figure 5 and 5 plots the movement of bond yields and the S&P500 in the crisis. All of these prices dropped at the same time. Every investor is "marginal" in all of these assets.

To be clear, I think the evidence is compelling that "small" arbitrage opportunities in hard-to-trade markets during the fall of 2008 are linked to intermediary problems. I put "small" in quotes, because an economically small arbitrage opportunity – say, a 1% deviation from covered interest parity – while not enough to attract long-only interest on one side or the other, represents a potentially enormous profit for a highly leveraged arbitrageur. But a 1% deviation is still small from the perspective of the overall economy.

I have similar doubts about the view that business and consumer debt is the major driver of asset prices and macroeconomics, rather than relatively minor, if important, epicycles. If bad times mean that the consumer will be close to the default limit, then why borrow so much in the first place? Buffer stock models require very high discount rates to eliminate this natural tendency to save up enough assets to avoid the bankruptcy constraint, and though the average person may be constrained, the

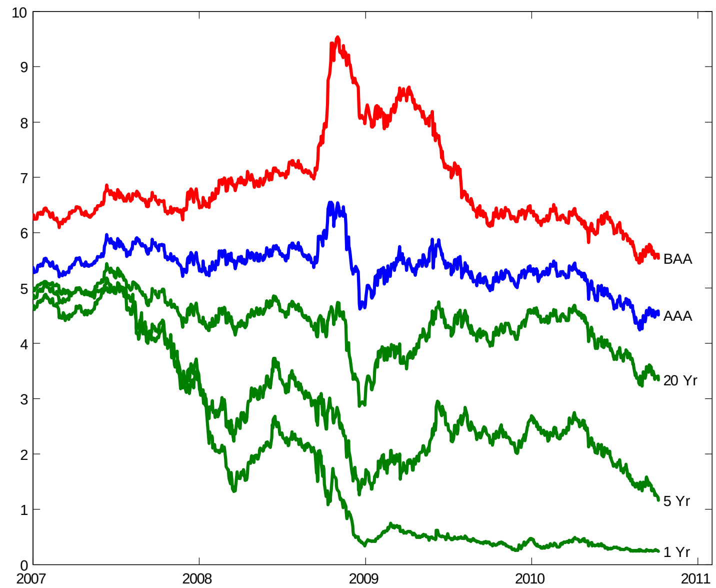


Figure 5: Bond yields

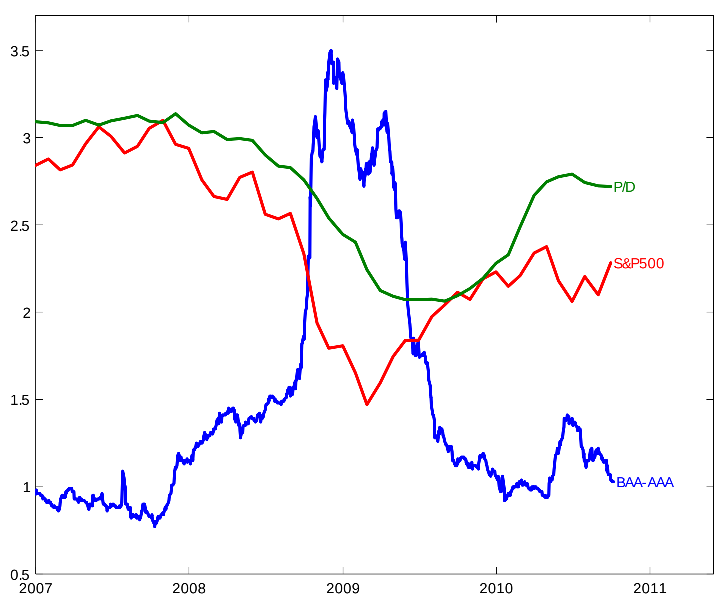


Figure 6: Bond yields and S&P 500

average dollar driving the risk-bearing capacity of the market is held by an unconstrained consumer.

The institutional finance view also does not easily explain why asset prices are so related to macroeconomic events. Losing money on intermediated and obscure securities is not naturally related to recessions. The 2007 hedge fund collapse did not lead to a recession.

One needs to imagine reverse causality, a new model of macroeconomics by which financial events spread to the real economy not vice versa. That's an exciting possibility, actually, and the core of the bustling frictions-based macro-finance research agenda. But at this stage it's really no more than a vision – models adduce frictions far beyond reality, such as that no agent can buy stocks directly, and data analysis of one event.

So, in my view, institutional finance and small arbitrages are surely frosting on a cake, needed to get a complete description of financial markets in times of crisis. But are they also the cake? And are they the meat and potatoes and vegetables of normal times, and the bulk of movements in broad market indices, and the explanation for their correlation with macroeconomics? Or can we understand the big picture of macro-finance without widespread frictions, and leave the frictions to understand the smaller puzzles, much as we conventionally leave the last 10 basis points to market microstructure, but do not feel that microstructure issues drive the large business cycle movements in broad indices?

Again, though, my main point is to point out the many commonalities, and only slightly to complain about differences. Theories based on debt deliver the same central idea, that the risk bearing capacity of the market declines in bad times.

The theories outlined so far differ only in the state variable for expected returns – consumption relative to recent values, news about long-run future consumption, cross-sectional risk, or leverage; balance sheets of individual consumers or those of leveraged intermediaries. All four state variables are highly correlated, and all four capture the idea that investors are scared of recessions.

### 4.5. Rare disasters

Robert Barro (2006) has recently taken up an idea of Thomas Reitz (1988), that the equity premium and other asset pricing phenomena can be understood with rare disasters. With Barro's inspiration, this idea has expanded substantially.

Looking back at the basic asset pricing equation,

$$E_t(R_{t+1}) - R_t^f = cov_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}, R_{t+1}\right] \leq \sigma_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]\sigma_t\left(R_{t+1}\right).$$

If people worry about rare events with very low consumption growth, then the variance of marginal utility in investors heads will be larger than the variance of consumption growth that we measure in a sample that doesn't include the rare event.

The basic idea is reasonable; that people worry about severe events when buying securities. People in California still worry about large earthquakes, though we haven't seen one since 1906, and rare events are priced in to earthquake insurance. More generally the fact that we are speaking English and worrying about the US equity premium, not German, Russian or Chinese represents some luck of the sample. Over the last century some truly disastrous events have occurred, and they haven't been all that rare.

In the time-varying disasters view, risk aversion does not rise. Instead, a combination of consumption risk drives up equity premiums, and asset market risk also raises expected returns. Both terms on the right hand side of my equation rise. Barro might answer Shiller that low prices really do forecast low dividend growth, we just didn't see the rare events.

The quick objection is that we really should have seen more disasters if they are large or frequent enough to account quantitatively for the equity premium with low risk aversion. This observation has led to a huge data controversy over just how many disasters we have seen, in the US and abroad, how to define a disaster, and what it constitutes.

Dark matter is a deeper objection. Unobserved rare events are already to some extent a dark matter assumption. But to get the central phenomena addressed by the habit model – return predictability, price/dividend ratio volatility, varying volatility, all of this correlated with business cycles – we need *time-varying* probabilities of rare disasters. That's really dark matter – unless one proposes some way of independently tying the time-varying probability of rare disasters to some data, which has not happened. One might surmount the dark-matter criticism if one assumption about time-varying disaster probability could reconcile multiple asset prices, but as Xavier Gabaix (2012) has pointed out has pointed out that, to make sense of the different asset classes, one needs to assume a asset-specific time-varying loading on the disaster risk.

Finally, the correlation of asset prices with business cycles relies on a correlation of business cycles with a time-varying disaster probability. As a correlation between short-run consumption growth and long run news is not totally implausible, neither is this correlation. But it is one more exogenous assumption, and one step harder to test than the correlation of consumption growth with long-run news.

In sum, the rare disaster view also requires a complex set of assumptions about the exogenous endowment process in order to explain the appearance of time-varying risk premia. Like the long-run risks model, measuring this process independently is challenging. The tie between observables and time-varying rare disaster probability is even harder to measure than the tie between observables and very long run consumption growth.

## 4.6. Probability assessments

Another class of models generalizes rational expectations. Suppose people's probability assessments are wrong. I include the bulk of behavioral finance here, which uses survey, psychology, and lab experiments to motivate wrong probability assessments, as well as modifications of preferences under the labels "Knightian uncertainty," "ambiguity aversion," and "robust control," which Lars Hansen and Tom Sargent have written about influentially.

The basic asset pricing equation, with the expectation written as a sum over states $s$, is

$$p_0 u'(c_0) = \beta \sum_s \pi_s u'(c_s) x_s$$

where $p_0$ is time zero price, $s$ indexes states of nature at time $1$, and $x_s$ is a payoff. (Typically $x_s = d_s + p_s$ will include a dividend and tomorrow's price.)

As this equation emphasizes, *probability and marginal utility always enter together.* There is no way to tell risk aversion – marginal utility – from a probability distortion from price $p$ and payoff $x$ data alone. That is, there is no way to do it without some restriction – some model that ties either probability distortions or marginal utility to observable data. This statement is just the modern form of Fama's "joint hypothesis theorem" that you can't test efficiency ($\pi$) without specifying a model of market equilibrium ($u'(c)$). Likewise, absent arbitrage opportunities, there is always a "rational" model, a specification of $u'(c)$ that can rationalize any data.

Given these facts, one would have thought that arguments over "rational" vs. "irrational" pricing, using only price and payoff data, would have ended the minute Fama's (1970) essay and joint hypothesis theorem were published. They have not, and to this day half of the published papers in finance claim to find one or another resolution to this argument without tying probabilities or marginal utility to data in some way.

The solution, of course is to tie either probabilities or marginal utility to observable data, in some rejectable way. In our general formula, if $\pi_s(Y)$, where $Y$ is measurable, then it becomes a testable theory. Behavioral economists have resisted tying themselves down in this way.

Without such a specification, "sentiment" is another dark-matter ex-post explanation. However, time-varying rare-disaster probabilities, not separately measured, or time-varying news about far-future incomes, not separately measured, are as much dark matter and really can't throw stones here.

The robust business cycle correlation of price ratios, explained by waves of "optimism" and "pessimism," is another troublesome fact. The habit model, by reverse engineering, captures this fact. People accept more risk in good times, and are resistant to accept the same risk in bad times. A model of probability mistakes has to explain why people are irrationally optimistic in booms and irrationally pessimistic in recessions. Again, that's not impossible, but it remains on the agenda for future research. I think the most natural explanation is reverse-causation, that asset price "bubbles" and "busts" affect the macroeconomy. But such a macroeconomic model has yet to be written down.

Behavioral economists point to surveys, in which people report amazing possibilities as their "expectation." But jumping from "what do you expect" in a survey to "what is your true-measure conditional mean" in a model is a big jump.

The survey never asks "by the way, did you report your risk-neutral or true-measure mean?" They don't ask that for obvious reasons – people would look at the questioner with dumbfounded disbelief. But the question is crucial. The risk-neutral probability is the actual probability times marginal utility,

$$\pi_s^* = \pi_s \beta \frac{u'(c_s)}{u'(c_0)} R^f.$$

With risk-neutral probabilities, price is the expected payoff, discounted at the risk free rate.

$$p_0 = \frac{1}{R^f} \sum_s \pi_s^* x_s = \frac{1}{R^f} E^*(x)$$

Now, imagine that prices are absurdly high, true expected returns are extremely low, you ask in a survey what investors "expect," and they answer that they "expect" good returns (good expected $x$) in the future, justifying the price. Irrationality confirmed! But without the followup question, if respondents reported the *risk-neutral* probabilities, they are not being irrational at all. The price *is* the risk-neutral expectation of payoff!. So the question "are those true-measure or risk neutral probabilities?" is not a technicality, it's the entire issue.

And it would be entirely sensible for people to think about and report risk-neutral, not true probabilities. Since probability and marginal utility always enter together, risk-neutral probabilities are a good sufficient statistic to make decisions. Risk neutral probabilities mix "how likely is the event?" with "how much will it hurt if it happens?" That combination is really what matters. Avoid stubbing your toe on the door jamb, yes. But put more effort into avoiding getting run over by a truck – though it's much less probable, it hurts a lot more.

More generally, the colloquial word "expect" is centuries older than the mathematical concept of true-measure conditional mean. Statisticians borrowed a colloquial word to describe their concept. But unless trained in statistics or economics (and, as teachers will ruefully note, actually remembering anything from that training) there is no reason to believe that a surveyed person has the statical definition in mind rather than the colloquial definition.

The online Oxford English Dictionary defines "expect" as to "regard (something) as likely to happen," and does not mention the statistical definition. So even a literate person does not know you're asking for the conditional mean. The online etymology dictionary cites the use of "expect" in something like the modern sense, "regard as about to happen," from the 1600s. Its Latin root, *expectare*, to "await, look out for, desire, hope, long for, anticipate, look for with anticipation" goes back further.

The distinction between risk neutral and real probabilities was formalized in 1979 by Harrison and Kreps. Whether the average survey respondent knows it today is a good question. The OED's lovely quotation, "England expects that every man will do his duty", Lord Nelson at Trafalgar, sounds behaviorally optimistic as an expression of conditional mean.

The ambiguity aversion literature also distorts probabilities. For reference I'll write down a heuristic equation,

$$p_0 u'(c_0) = \beta \sum_s \pi_s u'(c_s) x_s$$
$$\{\pi_s\} = \arg \min_{\{\pi \in \Theta\}} \max_{\{c\}} \sum_s \pi_s u(c_s)$$

The probabilities are chosen, in a restricted set, as those that minimize the maximum attainable utility. The investor focuses on the worst-case scenario in a set, and devotes all his attention to that case.

Obviously, hard questions remain. Most of all, just what is the restricted set $\Theta$? If you worry about meteorites falling from the sky, maybe you should worry about anvils and pianos too? Again, also, tying the distorted probabilities to measurable data remains the key to understanding variation in prices over time.

## 4.7. Summary

I have let myself stray too far to the realm of grumpy old guy who wants to defend his habits. The real picture is that many ideas give about the same result. There is an extra, recession-related state variable,

$$M_{t+1} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} Y_{t+1}$$

and the tendency for assets to fall when $Y_t$ is bad drives risk premiums, and changes in the conditional density of $Y$ drive time-varying risk premiums. Each of the models suggests different candidates for $Y_t$. But these candidates are are highly correlated with each other in the data, and sensibly indicative of fear or bad times. This fundamental unity is worth building on in the future, and I suspect models of the future will include elements from several of these insights, as well as (don't underestimate this!) analytical tractability to representing the common ideas in a nice quantitative parable.

The extent of my grumpy old guy comments are just to point out that, despite the relative popularity of the newer models, no model yet decisively improves on habit in describing the equity premium/risk free rate puzzles, and more importantly time-varying, business-cycle related risk premia; return predictability; "excess" volatility; "bubbles" associated with business cycles, and the long-run equity premium. At least habits should still be in the running.

Moreover, I still score the habit model as doing well based on number of assumptions relative to predictions. The time-varying risk aversion at the center of the model is endogenous, and a simple function of consumption relative to its recent past. Most other models require carefully calibrated and complex exogenous driving processes, which in many cases (long run risks, rare disasters) are nearly invisible in the data, or approach vacuousness and ex-post storytelling, such as labeling a market rise a rise in "sentiment" or "selling pressure," without independent measurement. But these are challenges which the other approaches may well surmount. Again, my main point is that habit models remain analytically tractable and at least not deep in the frontier.

## 5. Risk-averse recessions

And now, let us glimpse the ghost of habit future.

It is time to unite these models that explain asset prices, with production, general equilibrium and macroeconomics. It is also time for asset pricing to bring its biggest lesson to macroeconomics. *Asset price fluctuations are all about variation in risk premiums, not variation in interest rates.* Asset price fluctuations are highly correlated with recessions. It follows, I think, that *recessions are all about varying risk* *premiums, not about interest rates and intertemporal substitution.*

Granted, merging macroeconomics and asset pricing is the rallying cry of the institutional finance / frictions research agenda. But, following on with habits and the many similar approaches laid out above in which relatively frictionless models can address the asset pricing phenomena – including the crisis, as emphasized by Figure 3– I'd like to speculate about the lessons of habit models and their relatives, in which time varying risk premiums pervade the economy, not just segmented financial markets, for macroeconomics.

Habits are common in macroeconomics, but usually in a one-period form, $(c_t - \theta c_{t-1})$ with a small value of $\theta$ such as 0.4. These preferences help to give a hump-shaped impulse-response function , inducing the kind of consumption-growth smoothing that we deliberately sought to ignore. The low value of $\theta$ and loglinearization of the model mean the risk aversion channel we emphasize is largely absent.

But integration of ether habits or similar asset pricing models with macroeconomics, to further illuminate both asset pricing and macroeconomics, is already a headily active branch of research. As examples, one need look no fruther than this conference. All of the keynote speeches have been broadly on this theme.

Martin Lettau, (Lettau Ludvigson and Ma 2015) presented a model in which the capital share is the central variable for asset pricing. Leonid Kogan (Kogan, Papanikolaou, and Stoffman 2015) integrated asset pricing with technological innovation, growth, and the birth and death of small firms. John Cambpell (Campbell, Pflueger and Viceira 2015) presented a sophisticated model combining habits and a new-Keynesian macro model to describe variation in nominal bond betas. Olesya Grishchenko (Grishchenko , Song, and Zhou 2015) presented another long-run risks model with inflation non-neutrality to address the term structure.

Other examples merging asset pricing and macroeconomics abound. For example, Adrien Verdelhan (2010) showed how two habit economies living side by side produce the forecastability of currency returns; the low interest rate country has higher risk premiums Lopez, Lopez-Salido, and Vazquez-Grande (2015) use slow-moving habits, extended to the utility of leisure, production using capital and labor, investment with adjustment costs, and Calvo-style price rigidities, to address the term structure of risk premiums.

Still, to my mind, this work largely incorporates important macroeconomic modeling ingredients to understand asset prices at a deeper level. I think the next step is to turn the invasion around and use the finance ingredients to understand macroeconomics at a deeper level.

At this conference, Anthony Diercks (Diercks 2015) presented a sophisticated new-Keynesian macro model, including long run risks to incorporate asset pricing facts, to address the optimal target inflation rate for monetary policy. As another example, DePaoli and Zabczyk (2012) construct a new-Keynesian model with external habits and a strong precautionary saving motive to discuss cyclical monetary policy. They find that precautionary saving or its disappearance means that policy should be more restrictive following positive productivity shocks, a common intuition.

But I hope we can go much further, and construct a full model of business cycles in which changes in risk aversion or risk bearing capacity are at the heart of the whole phenomenon of business cycles.

In traditional Keynesian models, recessions are about static flows. Consumption is a marginal propensity times income, $C = a + mpcY$; investment is a static function of interest rates $I = \bar{I} - br$, output is $Y = C + I + G + NX$, and so forth. Alas, intertemporal economics dethroned this approach as an economic model.

In the new-Keynesian models that dominate current macroeconomics, recessions are about intertemporal substitution. The key equation (as in real business cycle models) is

$$c_t = E_t c_{t+1} - \sigma r_t + \varepsilon_t^d$$

which is a loglinearization of our standard first order condition with a preference shock. Consumption is low when real interest rates are high because people shift consumption in to the future. In words, recessions are times when everybody is trying to save too much and consume too little, and savings is about trying to consume too much in the future.

But 2008 was not a time at which people became thrifty, saving more for a better tomorrow. In 2008, people stopped consuming and investing because they were scared to death. "The" interest rate on Treasuries at the center of conventional models – which fell, not rose – is the least interesting asset price in 2008. The stunning and coordinated risk in *risk premiums,* completely absent in most macro models – the spike in credit spreads, the collapse in stocks, the arbitrage opportunities in derivatives – was the central price phenomenon of the recession. Investment did not fall because interest rates rose. They didn't in this case, and there is almost no correlation between investment and interest rates in the data. Investment sensibly fell because (among other things) the interest rates on corporate bonds, and the yields on equities, the instruments actually used to finance investment, rose – all due to rising risk premiums – while interest rates declined. The correlation of investment with stock prices (Q) is excellent, both in booms and in busts, and through the financial crisis, as Figure 7 emphasizes.

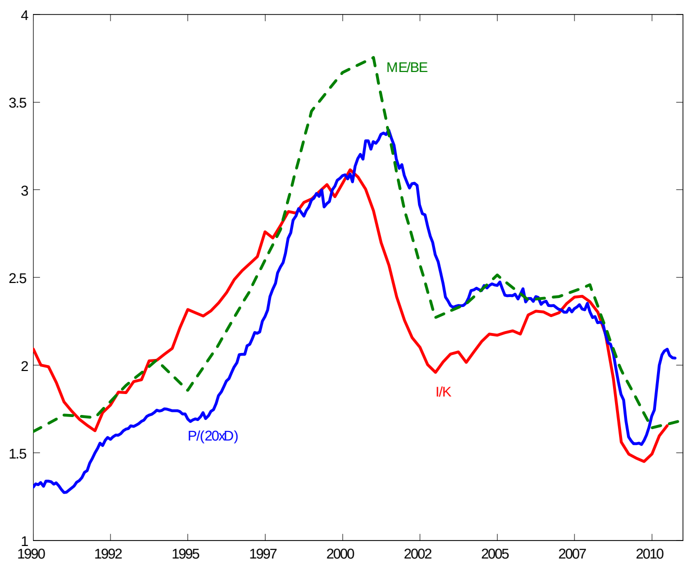


Figure 7: Investment to capital ratio, Market-to-book ratio, and price-dividend ratio.

My vision, then, goes something like this: A negative shock happens. It shouldn't really matter what the shock is, because we've never clearly seen underlying shocks that cause business cycles. I'll think of a small negative shock to wealth. Consumption falls a bit, and consumers get more risk averse. As they get more risk averse, precautionary savings rise, and consumption demand falls further. Price-dividend ratios fall, as in the endowment-economy habit model. Then investment falls as well, due to Q theory as illustrated in Figure 7. Though people want to save more due to precautionary savings, they want to save it in safe assets, not the risky opportunities offered by available technologies. Demand for government bonds rises, which also depresses inflation (there is a bit of fiscal theory of the price level in the latter channel). A decline in consumption, investment, and flight to quality pretty much define a recession.

The continuous-time equation for the interest rate is a good place to start fleshing out this vision. With a habit $x$, we have

$$r = \delta + \gamma \left( \frac{c-x}{c} \right) E\left( \frac{dc}{c} \right) - \frac{1}{2}\gamma(\gamma+1)\left( \frac{c-x}{c} \right)^2 \sigma^2.$$

As $c$ starts to fall, risk aversion starts to rise, and the last precautionary savings term rises. Fixing the interest rate, (set by the Fed, by foreign investment, by storage, or otherwise by technology), expected consumption growth $E(dc/c)$ has to rise. For expected consumption growth to rise, the level of consumption has to fall, (this is the standard new-Keynesian aggregate demand mechanism by which higher rates lower consumption) which raises risk aversion even more.

In standard models, (both new-Keynesian and real business cycle) the habit term is absent, and $\gamma$ is small. Since $\sigma$ (not $\sigma^2$) is of the same order as $E(dc/c)$, the second term on the right is unimportant. With habits or high risk aversion (needed so far in any model to account for the equity premium), the second term is all important. Squaring large risk aversion overcomes squaring small standard deviation. The big news from asset pricing for macro is, “don’t ignore precautionary savings!”

Many macro modelers have approached the 2008 period following the financial crisis by supposing a $\delta$ preference shock, a sudden increase in patience. They acknowledge this is a short hand for some other feature of a more fully fleshed out model. A rise in precautionary savings, in the third term, is exactly such a feature, relative to a model that ignores that term.

This effort needs to escape the Tallarini (2000) separation theorem, which otherwise hangs as a Modigliani-Miller warning against the whole enterprise. (Lopez, Lopez-Salido, and Vazquez-Grande 2015 call the phenonmenon “macro-finance separation.” ) In many models, quantity dynamics are driven by intertemporal substitution, and asset prices are driven by risk aversion, and the two don’t mix. Raising risk aversion raises the equity premium and depresses asset prices, but has no effect on quantity dynamics. Hence, macro can happily proceed ignoring equity premiums, and finance can tack on higher risk aversion to model asset prices, knowing that these modifications don't substantially affect the underlying quantity dynamics.

The intuition for this result is clear and suggestively robust. Typical adjustment cost technologies typical of Q theory allow the consumer/investor to trade less consumption today for more consumption in the future, spread across states of nature by technology shocks. (If you want an example with equations, think of $c_1 = \theta_1 f(k_0 + y_0 - c_0)$ with $theta_1$ random and $f(\cdot)$ concave.) But the distribution of the technology shocks is given. There is nothing the consumer can do to make this opportunity less risky.

The program I outlined here is obviously completely at odds with that separation. So how do we avoid macro-finance separation? The last equation suggests that precautionary saving is an important first ingredient. With important precautionary saving effects, raising risk aversion does change intertemporal substitution and thus the desire to save and invest overall.

The second ingredient, I think, is to enrich the production technology so that consumer/investors can shape the riskiness of the technological opportunities they face. Frederico Belo (2010) and Urban Jermann (2013) have recently explored specifications of technology that allow such choices. But much less radical changes can achieve the same ends. Here, I specify two production technologies, a risky one and a less risky one. When risk aversion rises, people want to shift investment from risky to less risky, facing adjustment costs and irreversibilities. This desire has strong consequences for quantities. Macro-finance separation relies on one production technology, so it can be circumvented by this real-side portfolio allocation effect.

## 5.1. Consumption: A two-period example

To get further with this intuition, we need to study the response of consumption to wealth. We can't do that from the first order condition alone. For the purposes of a speech, I'll work out a simple two-period model that you should be able to follow instantly. This model also shows nicely how habits capture many of the kinds of behavior and intuition that are used to suggest other kinds of models.

There are two periods. The representative consumer has an initial endowment $e_0$ and a random time-1 endowment $e_1$. The endowment $e_1$ can take on one of two values. His problem is then

$$\max \frac{(c_0 - x)^{1-\gamma}}{1-\gamma} + \beta E\left[\frac{(c_1 - x)^{1-\gamma}}{1-\gamma}\right]$$
$$c_1 = (e_0 - c_0)\, R^f + e_1$$
$$e_1 = \{e_h, e_l\}\;\; pr(e_l) = \pi.$$

I specify $\beta = 1/R^f = 1$ to keep it simple. The solution results from the first order condition

$$(c_0 - x)^{-\gamma} = E\left[(c_1 - x)^{-\gamma}\right]$$

i.e.,

$$(c_0 - x)^{-\gamma} = \pi(e_0 - c_0 + e_l - x)^{-\gamma} + (1-\pi)(e_0 - c_0 + e_h - x)^{-\gamma}$$

I solve this equation numerically for $c_0$

Figure 8 presents consumption $c_0$ for $e_h = 2$, $e_l = 0.9$, $x = 1$, $\gamma = 2$ and $\pi = 1/100$. The case that one state is a rare disaster is not special. In a general case, the consumer starts to focus more and more on the worst-possible state as risk aversion rises. Therefore, the model with any other distribution and the same worst-possible state looks much like this one. It is a simplification, not a strange special case. I set the simulation up so that in the bad state, $e_l = 0.9 < x$. The consumer has to do something to make sure that consumption exceeds the habit in this bad state.

Starting from the right, when first-period income $e_0$ is abundant, the consumer follows standard permanent income advice. The slope of the line connecting initial endowment $e_0$ to consumption $c_0$ is about 1/2, as the consumer splits his large endowment $e_0$ between period 0 and the single additional period $1$.

As endowment $e_0$ declines, however, this behavior changes. For very low endowments $e_0 \approx 1$ relative to the nearly certain better future $e_h = 2$, the permanent income consumer would borrow to finance consumption in period 0. The habit consumer *reduces* consumption instead. As endowment $e_0$ declines towards $x = 1$, the marginal propensity to consume becomes nearly one. The consumer reduces consumption one for one with income.

Figure 9 presents marginal utility times probability, $u'(c_0) = (c_0 - x)^{-\gamma}$, and $\pi_i u'(c_i) = \pi_i(c_i - x)^{-\gamma}, i = h, l$. By the first order condition, the former is equal to the sum of the latter two. But which state of the world is the more important considera-

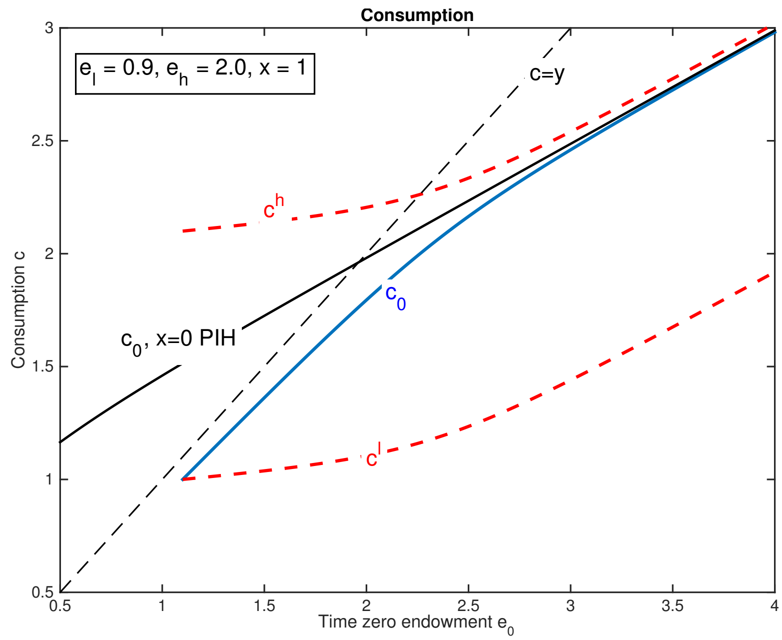


Figure 8: Conusmption as a function of time - 0 endowment

tion? When consumption is abundant in both periods on the right side of the graph, marginal utility $u'(c_0)$ is almost entirely equated to marginal utility in the 99 times more likely good state $(1-\pi)u'(c_h)$. So, the consumer basically ignores the bad state and acts like a perfect foresight or permanent-income intertemporal-substitution consumer, considering consumption today vs. consumption in the good state.

In bad times, however, on the left side of the graph, if the consumer thinks about leaving very little for the future, or even borrowing, consumption in the unlikely bad state approaches the habit. Now the marginal utility of the bad state starts to skyrocket compared to that of the good state. The consumer must leave some positive amount saved so that the bad state does not turn disastrous – even though he has a 99% chance of doubling his income in the next period ($e_h = 2$, $e_0 = 1$). Marginal utility at time $0$, $u'(c_0)$ now tracks $\pi_l u'(c_l)$ almost perfectly.

In these graphs, then, we see behavior that motivates and is captured by many different kinds of models:

- Consumption moves more with income in bad times.

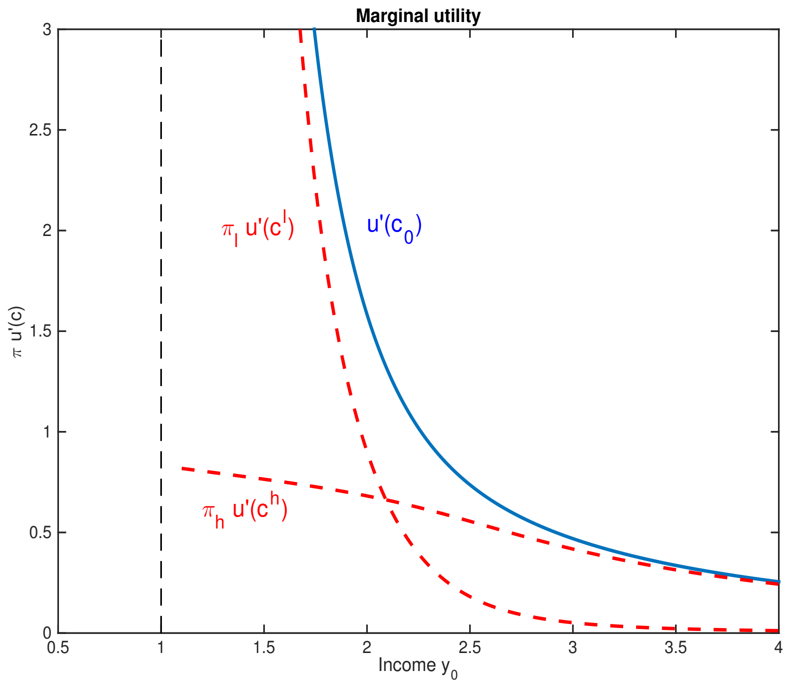


Figure 9: Marginal utility in the two-period habit model. Marginal utilities are weighted by the probability of each state

This behavior is familiar from buffer-stock models, in which agents wish to smooth intertemporally, but can't borrow when wealth is low. The habit view of this behavior has some differences, of course. The fundamental state variable is consumption relative to the recent past, not asset levels. That difference in view has some advantages: Buffer stock models have trouble confronting the fact that most consumers do have assets, which might be illiquid in the model but are pretty liquid in practice or on Craigslist. For this reason, high-income and high-wealth high-mpc consumers pose an even greater problem for buffer stock models.

- In bad times, consumers start to pay inordinate attention to rare bad states of nature.

This behavior is similar to time-varying rare disaster probability models, behavioral models, or to minimax ambiguity aversion models. At low values of consumption, the consumer's *entire* behavior $c_0$ is driven by the tradeoff between consumption today $c_0$ and consumption in a state $c_l$ that has a 1/100 probability of occurrence, ignoring the state with 99/100 probability. Here, it is not an irrational or ambiguity-averse assessment of that small probability which matters, it is the high marginal utility associated with that low probability, the necessity to keep consumption above habit no matter what. To slightly misquote Johnson, "Depend upon it, Sir, when a man thinks there is a 1/100 probability that he is to be hanged in a fortnight, it concentrates his mind wonderfully."

This little habit model also gives a natural account of endogenous *time-varying* attention to rare events. Creating any dynamics in an asset pricing or macro model requires such time-variation, which must be exogenous in standard rare events models. Here, bad times today (low $e_0$) lead the consumer to focus on the rare event.

Again, the point is not to argue that habit models persuasively dominate the others. The point is just that there seems to be a range of behavior that theorists intuit, and that many models capture. Habit models can and do produce the same behavior motivating the other models. And vice versa.

In bad times, risk aversion increases, risky asset prices fall, and risk premiums rise. That's the point of our original habit model. The price of a consumption claim is

$$p(c) = E\left(\frac{(c_1 - x)^{-\gamma}}{(c_0 - x)^{-\gamma}} c_1\right).$$

Figures 10 and 11 present this price and expected return, respectively. I contrast the price of the consumption claim with its riskfree valuation $E(c_1)$ (recall $R^f = 1$). I also compute the expected return $E(c_1/p(c_1))$ which I contrast with $R^f = 1$.

As you expect, the price of the consumption claim falls in bad times relative to the riskfree valuation. Correspondingly, there is a large rise in expected return relative to the riskfree rate. The riskfree rate is constant in this model, even though I do not have the nonlinear habit accumulation and delicate balance of intertemporal substitution and precautionary savings motives of our original habit model. Here, the linear capital accumulation technology enforces the riskfree rate. The unequal balance of precautionary savings and intertemporal substitution shows up by rising expected future consumption in Figure 10. In bad times, consumption is expected to grow faster than usual.

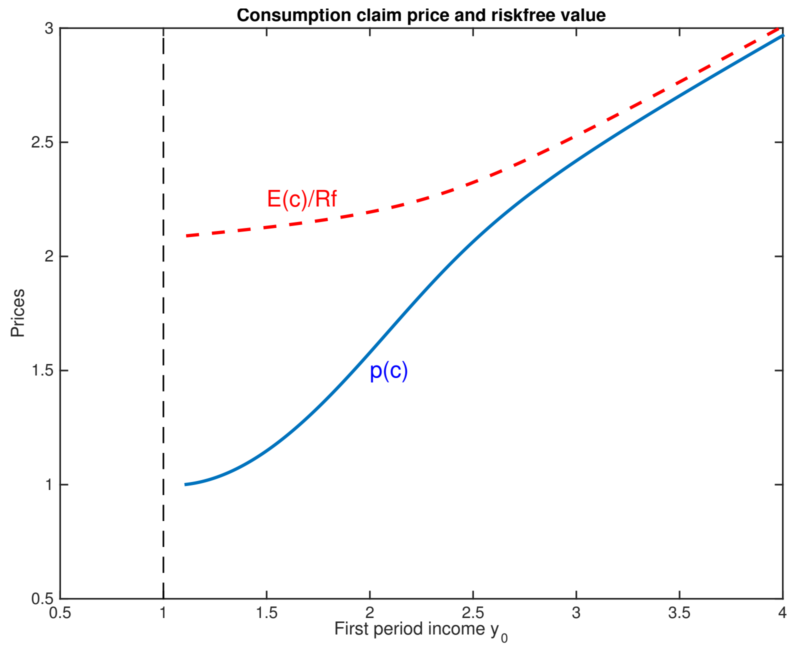


Figure 10: Price of consumption claim and riskfree value.

## 5.2. Investment

The plot of investment with stock prices, Figure 7, and the model's price-consumption ratio in Figure 10 would seem to seal the deal. Consumption "demand" falls, investment "demand" falls, and we have a recession. It's almost a multiplier-accelerator. However, getting such effects in a complete model is not as easy as it sounds. The next step is to add investment in a risky technology to the opportunity set.

By investing an amount $i_0$ at time zero, the consumer can get a random amount $\theta_1 i_0$ at time 1, where $\theta_1 = \{\theta_h, \theta_l\} = \{1.2, 0.9\}$. $\theta$ is the rate of return. In the good state and on average, with $R^f = 1$ or 0%, a 20% return is a very attractive opportunity. However, that attraction must be balanced with a 1% risk of a -10% return, coincident with a bad endowment shock.

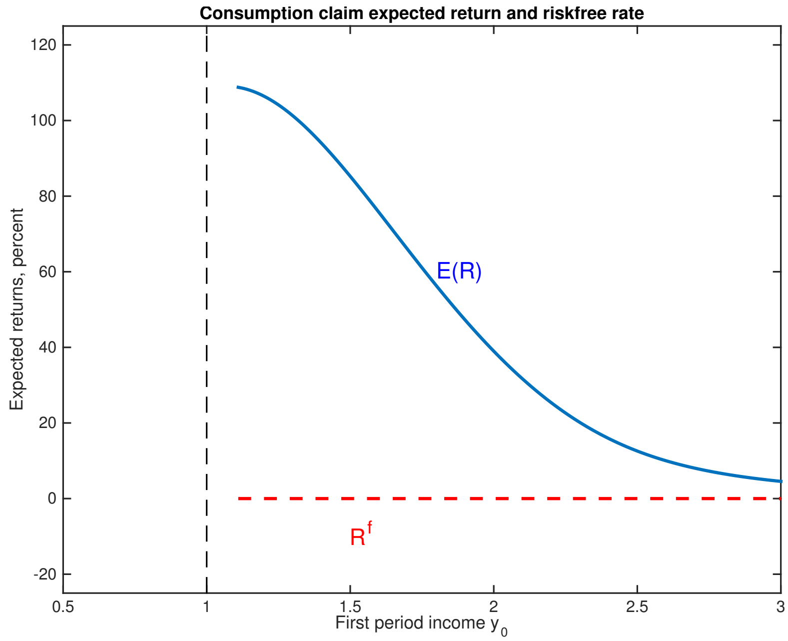


Figure 11: Expected return of consumption claim and riskfree rate

The technology is now

$$c_1 = e_1 + \theta_1 i_0 + B_0$$
$$c_0 = e_0 - i_0 - B_0/R^f$$
$$i_0 \geq 0;\ \theta_1 = \{\theta_h, \theta_l\}\,.$$

I impose positive investment at time zero. Without that feature, in bad $e_0$ times the consumer operates the production technology in a strongly negative manner, to shift consumption towards the low-productivity state. This ability is clearly unrealistic. A real model will have adjustment costs, irreversibility, and depreciation, and won't let you turn low productivity states into high ones by a negative capital stock. Here, $i_0 > 0$ does the same trick.

The consumer has *two* investment opportunities in this model. I want to think of $\theta$ as real, physical investment in productive but risky opportunities. I want to think of $B$ as riskfree storage, government bonds, or borrowing/lending abroad. Thus, we can capture the effects of risk aversion in shifting the *composition* of investment demand as well as overall savings/investment.

Maximizing the same objective

$$\max \frac{(c_0 - x)^{1-\gamma}}{1-\gamma} + \beta E \left[ \frac{(c_1 - x)^{1-\gamma}}{1-\gamma} \right]$$

with this enhanced technology, the model solution is characterized by two first order conditions

$$(c_0 - x)^{-\gamma} = E(c_1 - x)^{-\gamma}$$
$$(c_0 - x)^{-\gamma} = E\left[(c_1 - x)^{-\gamma}\theta_1\right] \text{ if } i_0 > 0.$$

Again, I solve numerically for $c_0$, using the same parameters.

Figure 12 presents consumption $c_0, c_h, c_l$, physical investment $i_0$, and riskfree debt $B$ for this model. $B$ is interchangeable with initial endowment $e_0$, so its level does not have any real significance. Figures 13 and 14 present the price and expected return of the consumption claim and the new risky technology.

The consumption behavior $c_0$ is quite similar to the previous case. When wealth $e_0$ is high, consumption follows the permanent income hypothesis, with a slope of about 1/2 in this two-period example. When consumption declines near habit, however, consumption tracks wealth almost one to one, as the consumer is concerned above all with keeping low-state consumption $c_l$ above habit.

The news here is the behavior of investment. For very low values of wealth $e_0$, investment is up against the zero constraint. The consumer would like very much to transfer consumption in to the bad state, by operating the technology in negative amounts. With that ruled out, he invests everything in the riskless opportunity.

Past wealth $e_0 \approx 2.2$, however, the attraction of the production technology's very high rate of return overwhelms the now lower risk aversion. As wealth increases, investment increases strongly. In this case, that increase is financed by borrowing. So we even see leverage increase in good times.

Conversely, we have the second part of the multiplier / accelerator story. As wealth decreases, equilibrium physical investment collapses, along with "deleveraging," reducing the debt used to finance that investment. As wealth decreases even

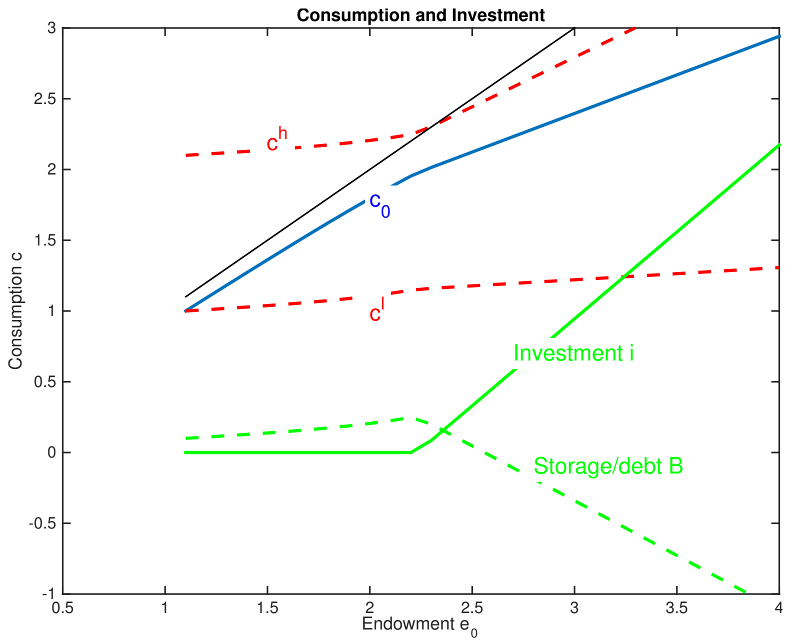


Figure 12: Consumption and investment as a function of time-0 endowment

more, investment collapses to zero and all savings go in to the riskless opportunity.

## 5.3. On to recessions (someday)

We have the two main ingredients of a theory of risk-averse recessions – consumption falls, with marginal propensities approaching one, and investment falls dramatically, along with leverage used to finance investment.

However, turning such "demand" into actual recessions requires additional steps, as always in macroeconomics. "Demand" may fall, but if $Y = F(K, L)$, why should output fall? Put another way, if the marginal utility of consumption rises so much, why not work harder to finance that consumption?

To illustrate the point in a simple static model, include labor hours $n$, less than total available hours $h$, and include leisure $h-n$ in the utility function. Let us also add the opportunity to produce the consumption good $c = wn$. Then, the consumer's objective is

$$\max\,(c-x)^{1-\gamma} + (h-n)^{1-\gamma} \;\text{ s.t.} c = wn$$

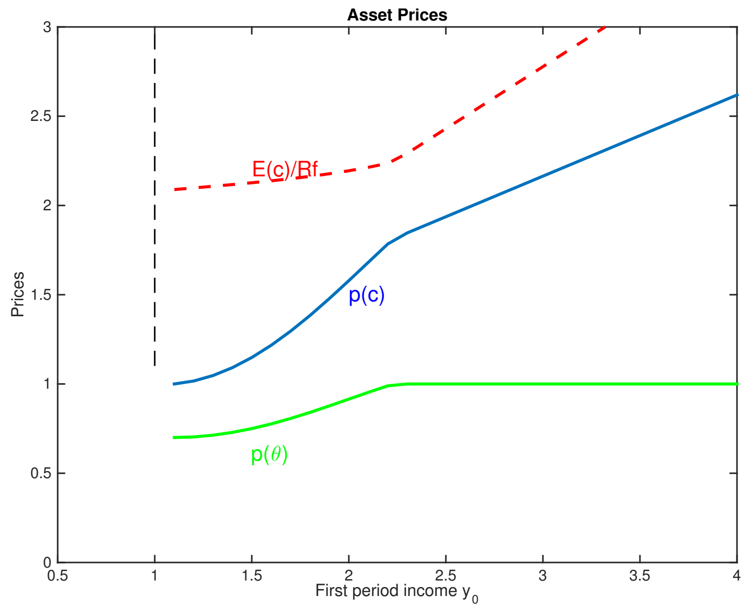


Figure 13: Price of consumption claim and investment technology

and the first order condition is

$$(c - x) = w(h - n).$$

So, in a state that consumption $c$ would otherwise fall close to habit $x$, the consumer will instead work more $n$, until labor hours rise towards the maximum available $h$. This example suggests that successfully incorporating leisure and other goods into the habit utility function will require habits of their own. And that is not unreasonable. Our ancestors worked 12 hours a day or more. Rearranging lives, and the fraction of a household that works, to accommodate much greater work hours might take time. The effective maximum number of hours may indeed evolve like a habit.

Similarly, the central puzzle of macroeconomics is the dissonance between saving and investment. If consumers want to save more, why does investment fall?

The usual response to these two puzzles is to add frictions, such as sticky prices and wages and monopolistic competition, such that output and labor effort follow consumption and investment "demand," not labor supply or the supply of savings.

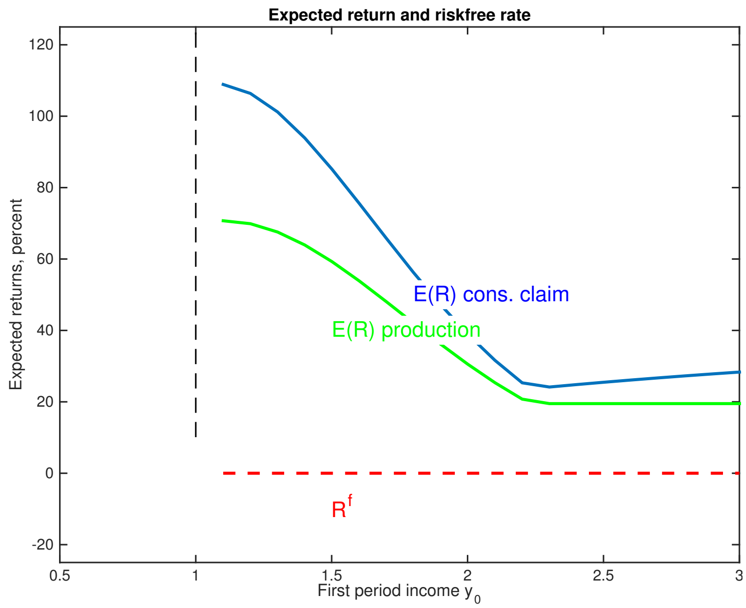


Figure 14: Expected return of consumption claim and investment technology.

One may end up following that traditional route. However, it strikes me that in this context, risk-averse recessions may emerge even without relying on nominal stickiness.

As we map the abstractions of the model into reality, two keys are important. First, let us think about riskless investment $B$ as storage, government debt, or international borrowing and lending. The interpretation as government debt I think is the clearest, as the 2008 recession featured a "flight to quality" surge in the demand for government debt. That labeling requires the slight of hand though that in the model, the government really can transfer real resources through time, while in reality government debt is a claim on future taxes. Only if the larger supply of government debt really is "invested," as our politicians love to call spending, in "infrastructure" or other projects that actually lead to larger future output and tax revenue, can we use the equations of this model with that label. But if we label it as such, then amounts invested in $B$ don't count to output or investment.

Then the dramatic portfolio *shift* in investment from the consumer's point of view, from risky technology to risk-free government debt, is a dramatic *reduction* in actual investment in to private capital stock.

Similarly, let us regard all private production as, to some extent, investment in a risky project. When a worker shows up at a car factory, steel factory, or even a bank, he or she is not producing a consumption good that can be consumed immediately. He or she is participating in a risky investment project. In reality, there really is nothing a typical worker can do, if the risky firm he or she is working at in Fall of 2008 shuts down, to produce anything of immediate consumption value. It even takes time to sign up to drive for Uber. The storied stockbrokers selling apples may not show a failure of the employment market, but the paucity of production opportunities to create consumption goods *today*.

As a set of equations that captures these ideas, let us write the model now as

$$\max(c_0 - x)^{1-\gamma} + E(c_1 - x)^{1-\gamma} s.t.$$

$$c_1 = e_1 + \theta_1 \min(i_0, n_0) + B_0$$
$$c_0 = e_0 - i_0 - B_0$$
$$i_0 \geq 0; h > n_0 > 0$$

In this formulation, labor $n_0$ falls exactly with investment $i_0 = n_0$. By specifying an inelastic labor supply, the solution of this model is exactly the same as the last model, so I don't have to solve any more equations. Now the fall in investment $i_0$ is the fall in labor $n_0$ and a fall in output.

Thus, by identifying the private economy as entirely and unavoidably devoted to the risky production technology, we have private output decline, private investment decline, and private labor decline in bad times, without the need for any stickiness.

The central mechanism is that which conventional macroeconomics rules out: because *risk aversion* increases, people want to *reallocate* investment, both of their resources and their labor effort (if they could) from risky to riskfree technologies. Private technologies are inherently risky, so we see the huge demand for government debt, and the collapse of private output, investment, and labor.

Of course this is only a suggestive and very stylized two-period model. The point, for an essay such as this: The big lesson of finance is that *risk premiums* vary over time, coordinated across asset classes, and correlated with recessions. Habit models capture that mechanism. The lesson of finance for macroeconomics then ought to be, that risk premiums and risk aversion, not riskfree rates and intertemporal substitution, are the central features of recessions. The fall in investment coincident with a rise in savings at the center of Keynesian economics can result from the fact that people want to *reallocate* investment to less risky projects even more than they want to save and invest more overall.

But all that awaits a real, complete, dynamic model.

## 6. Summary

In summary, we have learned that asset prices correspond to a large, time-varying, business-cycle correlated risk premium. This risk premium means that prices ratios and other variables forecast returns; not dividend growth, and thus that the risk premium accounts for the high volatility of price ratios.

A representative consumer model with habit preferences captures this phenomenon. It does so parsimoniously, in that the variation in risk premium is endogenous, and with a specific and rejectable independent measurement of its state variable, the relation of consumption to its recent past.

Lots of other modeling approaches capture the same facts, with a wide range of alternative underlying ideas and intuitions, including long-run risk, idiosyncratic risk, wealth shifts among agents with heterogenous preferences, debts and balance sheets, psychological or ambiguity-averse probability distortions, and time-varying rare disaster probabilities. None of these modeling approaches stands above the others in the list of facts so far addressed. A serious effort to distinguish them has not been made. But, given the fact that the state variables are so correlated, and that the models are all quantitative parables not detailed models-of-everything meant to be literally true, that effort may not be worth the bother. They differ, as I have pointed out, somewhat in the ratio of assumptions to predictions, and the amount of "dark matter" invoked to explain various phenomena, and more deeply they differ in the analytical convenience they each have in capturing the common ideas. The latter may be the most important feature for modeling developments.

As I look to the future, it seems time for this body of empirical and theoretical knowledge to invade macroeconomics. Recessions are phenomena of *risk premiums*, risk aversion, risk bearing capacity, desires to shift the composition of a portfolio from risky to risk free assets, a "flight to quality," not a phenomenon of intertemporal substitution, a desire to consume more tomorrow vs. today.

My vision applies equally if one thinks the variation in risk premium is "irrational," or the result of intermediary agency frictions. If one takes that view, then via the admirably fitting (see Figure 7) Q theory, financial market movements drive business cycles, and we're looking in all the wrong places for the causality (real to asset price vs. asset price to real) and nature of business cycles.

That invasion strikes me as even more interesting and productive than what we have accomplished so far, and most of my purpose in giving this talk has been to encourage you to join me in its quest.

I hope it won't take another 20 years!

# 7. References


Bansal, Ravi, and Amir Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles" *The Journal of Finance* 59, 1540-6261.

Bansal, Ravi, Dana Kiku, and Amir Yaron. 2012. "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices." *Critical Finance Review* 1, 183-221.

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

Beeler, Jason, and John Y. Campbell. 2012. "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment." *Critical Finance Review* 1, 141-182.

Belo, Frederico. 2010. "Production-Based Measures of Risk for Asset Pricing." *Journal of Monetary Economics* 57, 146-163.

Brunnermeier, Markus, 2009. "Deciphering the Liquidity and Credit Crunch 2007-08." *Journal of Economic Perspectives* 23, 77-100.

Campbell, John Y., and John H. Cochrane. 1999a. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy* 107, 205-251.

Campbell, John Y., and John H. Cochrane. 1999b. "By Force of Habit. Manuscript with Unpublished Appendices." http://faculty.chicagobooth.edu/john.cochrane/research/papers/habit.pdf

Campbell, John Y., Carolin Pflueger and Luis M. Viceira, 2015. ?Monetary Policy Drivers of Bond and Equity Risks.? Manuscript, Harvard Univeristy. http://scholar.harvard.edu/files/campbell/files/campbellpfluegerviceira_20150609.pdf?m=1433945372

Chen, Hui, Winston Dou, and Leonid Kogan. 2015. "Measuring the 'Dark Matter' in Asset Pricing Models." Manuscript, MIT Sloan School.

Cochrane, John H. 1988. "How Big is the Random Walk in GNP?" *Journal of Political Economy* 96, 893-920.

Cochrane, John H. and Monika Piazzesi. 2005. "Bond Risk Premia." *American Economic Review* 95, 138-160.

Cochrane, John H. and Monika Piazzesi. 2008 "Decomposing the Yield Curve." Manuscript, University of Chicago and Stanford University. http://faculty.chicagobooth.

edu/john.cochrane/research/papers/interest_rate_revised.pdf

Cochrane, John H., 2011. "Discount Rates: American Finance Association Presidential Address." *Journal of Finance* 66, 1047-1108. http://faculty.chicagobooth.edu/john.cochrane/research/papers/discount_rates_jf.pdf

Cochrane, John H., 2007. "Financial Markets and the Real Economy." In Rajnish Mehra, Ed. *Handbook of the Equity Premium* Elsevier. pp. 237-325. http://faculty.chicagobooth.edu/john.cochrane/research/papers/financial_and_real_proofs_aug_07.pdf

Constantinides, George M., and Darrell Duffie, 1996. "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy* 104, 219-240.

DePaoli, Bianca and Pawel Zabczyk. 2012. "Cyclical Risk Aversion, Precautionary Saving, and Monetary Policy." *Journal of Money, Credit and Banking* 45, 1-36. http://dx.doi.org/10.1111/j.1538-4616.2012.00560.x

Diercks, Anthony M., 2015. "The Equity Premium, Long-Run Risk, & Optimal Monetary Policy.". FEDS Working Paper No. 2015-087. SSRN: http://ssrn.com/abstract=2670122

Epstein, Larry G., Emmanuel Farhi, and Tomasz Strzalecki. 2014. "How Much Would You Pay to Resolve Long-Run Risk?" *American Economic Review* 104, 2680-97.

Epstein, Larry G., and Stanley E. Zin. 1989."Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57, 937-69. doi:10.2307/1913778.

Fama, Eugene F., 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *The Journal of Finance* 25, 383-417.

Gârleanu, Nicolae, and Stavros Panageas. 2015. "Young, Old, Conservative, and Bold: The Implications of Heterogeneity and Finite Lives for Asset Pricing." *Journal of Political Economy* 123, 670-685.

Gabaix, Xavier. 2012. ?Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance.? *Quarterly Journal of Economics* 127, 645-700.

Grishchenko, Olesya V., Song, Zhaogang and Zhou, Hao, 2015, "Term Structure of Interest Rates with Short-Run and Long-Run Risks" SSRN: http://ssrn.com/abstract=2663357orhttp://dx.doi.org/10.2139/ssrn.2663357

Hansen, Lars Peter and Ravi Jagannathan. 1991. "Implications of security market data for models of dynamic economies." *Journal of Political Economy* 99, 225-262.

Hansen, Lars Peter, John Heaton and Nan Li. 2008. "Consumption Strikes Back?: Measuring Long Run Risk." *Journal of Political Economy* 116, 260-302.

Hansen, Lars Peter and Thomas J. Sargent, 2001. "Robust Control and Model Uncertainty," *American Economic Review* 91, 60-66.

Harvey, Campbell R., Yan Liu and Heqing Zhu. 2016. "... and the cross-section of expected returns." Forthcoming, *Review of Financial Studies*. http://ssrn.com/abstract=2249314

Jagannathan, Ravi, and Yong Wang. 2007. "Lazy Investors, Discretionary Consumption, and the Cross-Section of Stock Returns." *The Journal of Finance* 62, 1623-1661.

Jermann, Urban J. 2013. "A Production-Based Model for the Term Structure." *Journal of Financial Economics* 109, 293?306.

Krishnamurthy, Arvind and Zhiguo He, 2013."Intermediary Asset Pricing?" *American Economic Review* 103.

Kogan, Leonid, Dimitris Papanikolaou, and Noah Stoffman, 2015, "Winners and Losers: Creative Destruction and the Stock Market." Manuscript, MIT. https://dl.dropboxusercontent.com/u/7804081/KPS_GE_2015.pdf

Lettau, Martin and Sydney Ludvigson. 2001. "Consumption, Aggregate Wealth and Expected Stock Returns," *Journal of Finance* 56, 815-849.

Lettau, Martin, Syndey C. Ludvigson, and Sai Ma. 2015. "Capital Share Risk and Shareholder Heterogeneity in U.S. Stock Pricing." Manuscript, University of California at Berkeley. urlhttp://www.sydneyludvigson.com/s/lbs.pdf

Lopez, Pierlauro and Lopez-Salido, David and Vazquez-Grande, Francisco, 2015. "Nominal Rigidities and the Term Structures of Equity and Bond Returns" SSRN: http://dx.doi.org/10.2139/ssrn.2626619

Oster, Emily, Ira Shoulson, and E. Ray Dorsey. 2013. "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease." *American Economic Review* 103, 804-30.

Piazzesi, Monika, Martin Schneider and Selale Tuzel. 2007. "Housing, Consumption, and Asset Pricing?" *Journal of Financial Economics* 83, 531-569.

Reitz, Thomas A. 1988. "The Equity Risk Premium: A Solution," *Journal of Monetary*

*Economics* 22, 117-31.

Schmidt, Lawrence. 2015. "Climbing and Falling Off the Ladder: Asset Pricing Implications of Labor Market Event Risk." Manuscript, University of Chicago.

Shiller, Robert J., 2014. "Speculative Asset Prices," *American Economic Review* 104, 1486-1517.

Shiller, Robert J., 1981. "Do stock prices move too much to be justified by subsequent changes in dividends?"' *American Economic Review*, 71, 421-436.

Tallarini, Thomas D., 2000. "Risk-Sensitive Real Business Cycles," *Journal of Monetary Economics*, 45, 507-532.

Tallarini, Thomas D., and Harold H. Zhang. 2005. "External Habit and the Cyclicality of Expected Stock Returns." *Journal of Business* 78, 1023-1048.

Verdelhan, Adrien. 2010. "A Habit-Based Explanation of the Exchange Rate Risk Premium." *Journal of Finance* 65, 123-145.

Wachter, Jessica. 2006. "A consumption-based model of the term structure of interest rates." *Journal of Financial Economics* 79, 365-399.