Comment on:
Are behavioral asset-pricing
models structural? ☆

Jessica A. Wachter*

Stern School of Business, New York University, New York, NY 10012, USA

1. Introduction

According to Stanley Zin, financial economists on both sides of the behavioral finance debate have failed to ask the right questions. To Zin, the most important question we should ask about behavioral models is whether they make accurate predictions. Structural models at least have prediction as their goal, whether or not they achieve it. Implicit in this view is the notion that structural models somehow "explain what is going on", that they can offer insights into why economic phenomena occur.

In most other fields of science, evaluating a structural model would not be a matter of debate. Because the goal of a structural model is prediction, the model would be evaluated directly using a controlled experiment. If in finance we could somehow change the parameters of our economy, or at least statistically detect such a change, we would also have a direct test of a structural model. The former is clearly impossible; the latter, Zin argues, is unreliable.

The question of how to evaluate a structural model without experiments is clearly a difficult one. Instead of attempting an answer, Zin argues that two commonly used criteria are wrong. One of these criteria, that a model appear realistic, looks good at first glance. But, quoting Friedman and Lucas, Zin argues that we are not trying for

---

*Tel.: +1-212-998-0799; fax: +1-212-995-4233.

E-mail address: jwachter@stern.nyu.edu (J.A. Wachter).

realistic models, we are trying for accurate ones. Ultimately models are better if they make better predictions, not because they are more realistic.[1]

This is a valuable point to keep in mind as one ponders the huge array of facts assembled by experimental psychologists over the years. Nonetheless, aficionados of behavioral finance might respond that allowing for behavioral assumptions does improve the models; added realism is merely a side benefit. Thus, we come to Zin's second false criterion: the fit to historical data.

That the fit to historical data can be a meaningless criterion is perhaps the major point of Zin's paper. Zin presents two examples of modeling strategies. The first allows complete freedom in the distribution of the aggregate endowment. The second allows complete freedom in investor's marginal utility (more precisely, the intertemporal marginal rate of substitution). Both models may appear to be structural, but in fact they are merely statistical representations of the data. The danger with behavioral models, according to Zin, is that they allow for state-dependent utility functions. Once marginal utility is state-dependent, it is possible to fit any moments of historical data if one allows for enough degrees of freedom. While the case Zin considers is extreme, it proves a point. When marginal utility is state-dependent, fitting the historical data is the wrong objective.

Ultimately, Zin argues, a subjective, non-sample based criterion will be needed to evaluate structural models. Somehow, asset pricing must come to a consensus on what models for the utility function are acceptable. To emphasize the subjective nature of the criterion, Zin refers to it as *reasonableness*. To those who say that establishing such a criterion is impossible, Zin points to the case of endowments. There is as much scope for spurious data-fitting for endowments as there is for preferences, and yet we do not see an explosion of endowment models purporting to explain the data. There appears to be a field-wide consensus on *reasonable* processes for endowments, why not for preferences?

In the conclusions, Zin suggests elements that might go into such a consensus. Principle among these is the requirement that utility be state-independent. This requirement would make spurious data fitting impossible, but I believe it goes too far. So what is the right answer? Below I argue that there is a relatively simple way to decide which models are *reasonable*. My criterion is objective and sample-based, and, best of all, it can be described in one word.

---

[1] I believe this issue is more complicated then Zin's initial discussion would suggest. Friedman and Lucas do not argue that realism is irrelevant, but rather that realism is not a goal to be pursued for its own sake. One would want a model to be robust to changes that make it more realistic; such changes should improve a model's ability to predict if it is on the right track. This does not necessarily rescue behavioral models, however. Because non-standard preferences do not aggregate as nicely as standard ones, some might argue that assuming non-standard preferences for the representative agent does not make the model any more realistic. Zin may have this argument in mind when he suggests, in the conclusion, that psychological evidence could be relevant, but only with an aggregation theorem.

2. Sir William's razor

I believe that parsimony lies at the root of what Zin refers to as reasonableness. A parsimonious model is a model in which the number of phenomena to be explained is much greater than the number of free parameters. Zin's examples seem so unreasonable to us because they are so terribly unparsimonious. The number of degrees of freedom *equals* the number of moments to be explained!

This idea is not new. In the early 14th century, Sir William of Occam was quoted as saying ''non sunt multiplicanda entia praeter necessitatem'', which the Encyclopedia Britannica translates as ''entities are not to be multiplied beyond necessity''. He must have said it often and sharply, because this principle came to be known as Occam's Razor. Later it received an Anglo-Saxon formulation as ''the principle of parsimony''. Galileo used this principle to establish that the earth goes around the sun, rather than the sun around the earth. Both theories perfectly fit the data, but one required what we might call multiple free parameters. Galileo's situation was similar to ours in at least one respect: he could not do experiments either. Instead, he successfully argued that when two models fit the data, the simplest one should win, no matter how much we would prefer to be at the center of the solar system.

Besides explaining what we do not like about Zin's example, the principle of parsimony also explains Zin's observation that ''the typical economist'' would not accept an explanation for the size of the equity premium based on the fifth moment. We would suspect that any model that relies on the fifth moment has also looked at moments one through four. Such a model could not possibly be a parsimonious explanation for the equity premium.

The principle of parsimony may be behind what Zin refers to as the consensus on endowments. I believe that Zin is correct in stating that this consensus does not entirely extend to preferences, even though preferences present at least as much of an opportunity for data-fitting mischief. This may be because adapting financial models to account for psychological principles is a new and potentially exciting idea. Enamored by psychological evidence, we may forget that adding bells and whistles to preferences creates noise and distraction, and ultimately could detract from a model's scientific value. This holds even if the bells and whistles are justified by psychological data. The goal of a structural asset pricing model is the parsimonious explanation of economic, not psychological, phenomena.

3. Reverse engineering revisited

Equipped with the principle of parsimony, it may be safe to let some models of state-dependent utility in the door. Zin's example illustrates the danger of reverse engineering. But reverse engineering is not entirely bad. Many behavioral models are reverse engineered to some degree, but this reverse engineering is entangled in assumptions motivated by psychological evidence. For this reason, it is helpful to

consider a model in which it is clear where the reverse engineering ends and the behavioral assumptions begin: the model of Campbell and Cochrane (1999).

The Campbell and Cochrane model has non-standard preferences, and so is included in Zin's list of behavioral models. Following standard convention, small letters denote natural logs; capital letters denote levels. Let C_t denotes aggregate consumption and M_t the stochastic discount factor. Campbell and Cochrane linearly add a state variable to $m_t = \log M_t$:

$$m_t = \log \delta - \gamma \Delta s_t - \gamma \Delta c_t.$$

This variable, s_t follows a process with linear conditional mean and non-linear variance, while consumption growth is i.i.d with constant mean:

$$\Delta s_{t+1} = (1 - \phi)(\bar{s} - s_t) + \lambda(s_t)\varepsilon_{t+1},$$
$$\Delta c_{t+1} = g + \varepsilon_{t+1}.$$

The variance term $\lambda(s_t)$ is reverse engineered to insure a constant riskfree rate. To see how this works, let r_0 be the constant riskfree rate. By the Euler equation

$$r_0 = -\log \delta + \gamma g + \gamma(1 - \phi)(\bar{s} - s_t) + \frac{\gamma^2 \sigma^2}{2}(\lambda(s_t) + 1)^2.$$

Therefore

$$\lambda(s_t) = \frac{\sqrt{2}}{\gamma \sigma}(-\log \delta + \gamma g + \gamma(1 - \phi)(\bar{s} - s_t) - r_0)^{\frac{1}{2}} - 1.$$

The resulting model has only four free parameters: ϕ, γ, δ, $\bar{s}$. Campbell and Cochrane choose ϕ and γ to match the first-order autocorrelation of the price-dividend ratio and the unconditional Sharpe ratio. That still leaves δ and $\bar{s}$. These are chosen so that, to a first approximation, the model is one of external habit formation. That is, $S_t = (C_t - X_t)/C_t$, where X_t is approximately a weighted average of past consumption. Certainly this model is reverse engineered, but the reverse engineering is limited to the riskfree rate and the two equity moments described above. The appeal of this model is based not on the fit to historical moments, but the fit to moments it was not designed to match. These include the equity premium, stock market volatility, and the volatility of the price-dividend ratio.

A potentially powerful way to evaluate this model and others like it, is to examine their implications for other asset classes. This is almost like running a controlled experiment. In Wachter (2001), I generalize the Campbell and Cochrane model so that interest rates are time-varying; the extra parameters are not regarded as free but are fitted directly from interest rate and consumption data. I find that the resulting model captures features of the yield curve and bond returns.[2] Success in these "out of sample tests" are arguably the best way to differentiate a structural model from a purely statistical one.

---

[2] In another example, Barberis et al. (2001) extend the static prospect-theory framework of Bernartzi and Thaler (1995) to capture time-varying stock returns. Barberis et al. require an additional behavioral assumption, known as the "house-money effect", which they motivate using psychological evidence. This assumption is not entirely uncontroversial, however. For a critique, see Brennan (2001).

4. Conclusion

Stanley Zin raises an import concern about behavioral models. In departing from rationality, behavioral models leave room for multiple degrees of freedom in the utility function. Taken to an extreme, this approach could reduce structural modeling to a tautological, data-fitting exercise. One might argue that psychological evidence itself restricts the parameters. There may be truth to this argument, but the wealth of (sometimes contradictory) psychological evidence, combined with the lack of training we have in evaluating this evidence, leaves it open to doubt.

Above, I argue that there is a solution: evaluate behavioral models on parsimony. Traditional structural models are implicitly parsimonious. Taking into account the degrees of freedom in a behavioral model puts it on the same footing as its traditional counterpart. In practice, how could this be accomplished? One way is to carefully count up the degrees of freedom of a model. Perhaps an easier way is to examine a model's implications for asset classes other than the one for which it was designed. Either way, Zin's question can be answered using the tools of econometrics, on which both sides of this debate still agree.

References

Barberis, N., Huang, M., Santos, T., 2001. Prospect theory and asset prices. Quarterly Journal of Economics 116, 1–53.

Bernartzi, S., Thaler, R., 1995. Myopic loss aversion and the equity premium puzzle. Quarterly Journal of Economics 110, 73–92.

Brennan, M.J., 2001. Discussion. Journal of Finance 56, 1292–1295.

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

Wachter, J.A., 2001. Habit formation and returns on bonds and stocks. New York University, working paper.