Review of the Equity Premium Puzzle

Benjamin Große-Rüschkamp*

Meet the Equity Premium Puzzle

The equity premium, the excess return of equity over relatively risk-free government bonds is well-documented and is of the magnitude of $3-8$ percent per year, depending on the period considered.[1] One basic explanation is that the equity premium is a risk premium that investors receive for bearing the additional risk associated with holding equity instead of bonds. This is the notion borne by Capital Asset Pricing Models. Because of theoretical shortcomings, macroeconomists have developed a generalization of the traditional CAPM. Consumption-based asset pricing allows for richer microfoundations and also takes into account the intertemporal dimension of portfolio investing (Breeden (1991)). So due to its theoretical foundations, consumption-based asset pricing seems to offer an excellent point of departure to pose the question about the "right" size of the 'risk premium'. Using the predominant macroeconomic asset-pricing model, Mehra and Prescott (1985) found that only a small fraction of the premium could, in fact, be accounted for by risk aversion under the neoclassical representative-agent paradigm. The finding, termed 'Equity Premium Puzzle' withstood scrutiny of

* Benjamin Große-Rüschkamp received his degree in Economics (B. Sc.) from the University of Bonn in 2011. The present article refers to his bachelor thesis submitted in September 2011.

[1] Mehra and Prescott (2008a) report for the annual percentage equity premium: 3.62 (1889−1933); 8.07 (1934 − 2005); 7.48 (1946 − 2005); Own estimations give 7.94 (1934 − 2010).

the profession, and puzzled macroeconomists came up with a host of new frameworks and expansions on existing concepts to reconcile the data with theory.

Mehra and Prescott (1985) set out to empirically test the implications of Lucas' seminal paper (Lucas (1978)) in which he was first to rigorously spell out the idea of relating asset price behaviour to consumption in the context of a complete market general equilibrium economy. The core of the modelled economy forms the representative agent. Here, the assumption of complete markets and isoelastic utility are sufficient for aggregation because in equilibrium, the marginal utilities of all agents are proportional (see, for example, Ljungqvist and Sargent (2004)). By deciding on a consumption path $\{c\}_{t=0}^{\infty}$, the representative agent maximizes the sum of its discounted expected utilities, i.e. $\max E_0\{\sum_{t=0}^{\infty} \beta^t U(c_t)\}$. The utility function $U(c, \alpha)$ is defined as $U(c, \alpha) = \frac{c^{1-\alpha}-1}{1-\alpha}$. Here, α represents the coefficient of relative risk aversion; at the same time, the inverse α^{-1} represents the coefficient of the elasticity of intertemporal substitution. This measures the agent's willingness to substitute consumption intertemporally. As both coefficients are time-invariant (i.e. constant), this class of utility functions is also referred to as CRRA/CIES[2] or isoelastic utility. Conveniently, CRRA is invariant to scale transformations so the demand for assets with risky payoffs is a linear function of wealth. As this makes it independent of initial distribution of wealth or endowment, it allows for aggregation of utility and the use of a representative agent. However, it also implies that a change of the preferences that affect the attitude towards smoothing consumption across time also affects the attitude towards smoothing consumption across states. This assumption may be problematic and is revisited later on. In equilibrium, the following relationship on asset prices should hold: $P_t U'(c_t) = \beta E_t\{U'(c_{t+1})(P_{t+1} + d_{t+1})\}$. This stochastic Euler equation, as presented by Lucas (1978) (equation (6)) has a convenient

[2] CRRA: Constant Relative Risk Aversion; CIES: Constant Intertemporal Elasticity of Substitution

interpretation. It expresses the intertemporal choice problem of the agent at time t. After performing some algebra we obtain what is known as 'Lucas asset pricing formula', namely $1 = E_t[m_{t+1}q_{t+1}]$, with $m_{t+1} = \beta\frac{u'(c_{t+1})}{u'(c_t)}$, where $q_{t+1} = 1 + R_{t+1}$ is the period gross return and m_{t+1} is a random variable and referred to as Stochastic Discount Factor. Performing some further operations, Mehra and Prescott calibrate the model. Using macroeconomic data, the free parameters remaining are β and α. Then, using values for parameters estimated by other economists,[3] Mehra and Prescott's 1985 estimation results in terms of admissible pairs of R^f and R^e that can be plotted using those values. Clearly off by more than an order of magnitude, the largest attainable value for the equity premium $R^e - R^f$ is 0.35, compared to 6.18 for the average estimated premium. On top, this value for the equity premium comes at the cost of a counterfactual average risk free rate of about 4 percent. This result represents the puzzle.

Preference-based solutions: Generalized Expected Utility and Habit Formation

The CRRA preference possesses many desirable traits, however, it links risk aversion and preference of consumption stability across time and across different states of nature. There is no compelling economic reason for why preferences concerning consumption patterns across time and states would have to be linked. It seems clear that increasing the risk-aversion of households by increasing the parameter value α of the original utility function will drive up the equity premium. Equity returns are more volatile, and thus perceived as more risky; so equity returns need to be high relative to risk-free returns to induce households to invest in them. This is even more true when increasing the value for α. So then,

[3] Friend and Blume (1975): $\alpha \geq 2$; Tobin and Dolde (1971) $\alpha = 1$; applying this information liberally, Mehra and Prescott (1985) restrict $\alpha \geq 10$.

why can we not just set α high and generate an equity premium that matches the observed? The issue is the following: by setting a high α, it is possible to produce a large enough equity premium through an increase of the household's risk aversion. At the same time, this also increases the household's aversion to unstable consumption paths over time. Given that consumption grows over time, there is little incentive to save. Rather, there is an urge to borrow against future income to increase consumption today. In the aggregate, however, households cannot borrow. What the inclination to borrow will do, though, is drive up returns of securities. Therefore, increasing α will raise returns; in particular the risk-less rate would be far higher than observed. Weil (1989) tests the asset-pricing implications in an economic environment otherwise identical to Mehra and Prescott (1985), but employing a GEU[4] preference ordering. A basic form of the aggregator used by Weil is the following: $U_t = \{c_t^{1-\rho} + \beta(E_t[U_{t+1}^{1-\gamma}])^{\frac{1-\rho}{1-\gamma}}\}^{\frac{1}{1-\rho}}$. This isoelastic utility function visibly disentangles the coefficient of risk aversion γ from the coefficient of intertemporal substitution, ρ. While intuition seems to suggest that a recalibration of preferences would significantly mitigate the Equity Premium Puzzle—adding a variable increases the degree of freedom for varying parameters—Weil (1989) demonstrates that this is not the case at all, and names the result of the counterfactual high rate, the 'Risk-free Rate Puzzle'. Disentangling the coefficients governing the attitude towards risk and intertemporal substitution does not remedy the familiar issues because the role of the coefficient of intertemporal substitution is actually strengthened by the separation: In a world of growing consumption, agents will once again have to be offered a counterfactually high return to be induced to postpone consumption.

Another potential way to solve the equity premium puzzle is to leave behind the assumption of per-period consumption as the only utility enhancing

[4] GEU: Generalized Expected Utility

measure, an approach grounded to some extend in the behavioural literature. An interesting case of attempts to resolve the puzzle is the use of models with habit formation: instantaneous utility is not simply derived from the per-period consumption-level, but is rather determined by relating it to the agent's history of consumption which may form an internal benchmark, also known as a habitual level of consumption ('internally measured habit'). In another variant, the agent derives utility from relative status, so his consumption is contrasted to the consumption-level of others ('externally measured habit'). In the first case, utility decreases if current consumption falls to less than a habitual level. The second case postulates a social dimension: there is a decline in utility if his consumption moves negatively relative to what the peers consume. This fall in utility is reasoned to be due to the perceived failure to 'keep up with the Joneses', as the old American idiom suggests. Conversely, in all cases, raising consumption above a habitual or 'Jones' level is accompanied by an increase in utility. How may this translate into a higher equity premium? One reason for the low premium in the original representative agent model is the low volatility in consumption. Such low volatility coupled with high returns on risky assets implies a very high risk aversion for agents. If, however, utility is not derived from consumption, but rather from ratios in consumption (relating current consumption to one of the above-mentioned reference levels), then even small variations in consumption may translate into more volatile, habit-adjusted levels. While there are a number of results in the literature that seem to successfully rationalize the data,[5] issues nevertheless remain. For one, in the context of habit-formation, it is possible for agents to experience welfare gains by a one-time (or periodic) lowering of the consumption level as subsequent gains ('consumption bunching') relative to that new lower habitual level would lead to increasing utility and thus quickly make

[5] Abel (1990), Constantinides (1990), Campbell and Cochrane (1999).

up for the utility lost in the period of the fall (Ljungqvist and Uhlig (1999)). In addition, these preferences can seem somewhat ad-hoc as they have no axiomatic foundations.

Market-based Solution

A promising feature of market incompleteness focuses on borrowing constraints in connection with viewing asset pricing through the lenses of a life-cycle model, as proposed by Constantinides, Donaldson, and Mehra (2002). To obtain new insights into the ongoing debate of the equity premium puzzle, they employ an overlapping generation model. New generations are born each period, which guarantees an infinite time horizon for the economy as a whole. In the first period, agents accumulate human capital while receiving low endowment income. The second period sees agents receiving a high, though stochastic wage income, whereas towards the end of their lives, in the third period, they retire and live off the wealth accumulated during the second period. The incomplete market features introduced by this model are such that no one can trade with yet unborn generations, nor can the young generation borrow against their future income, as they have no means of providing adequate collateral to secure their credit. The novelty of the approach arises from the implications derived by looking at the optimal portfolio held by each generation: The young face a situation where future wage and equity income are only weakly correlated (This is a key assumption of the model. Empirical support to the assertion is provided by Davis and Willen (2000)); investing in equity would thus be a good choice. However, not only is their income low, but with consumption smoothing on their mind, they are also extremely hesitant to spend their period-one income on anything else than consumption goods. Through the borrowing constraint, non-participation in the securities markets by the agent of the young generation arises endogenously. The middle-aged, on the contrary, have their wage uncertainties resolved, but are looking to save for old age and retirement. When they are old, they will have to live off their assets. Then, their consumption will be strongly correlated to the return of their assets. Hence they prefer to invest in assets with low variance and will buy mostly risk-free bonds. Given that the middle-aged fraction of the population provides the marginal investor, this setup provides a plausible justification for both a low return on risk-free assets as well as a very high return on equity and thus, a high equity premium. While there are still some issues such as relaxing the no-bequest assumption to be satisfyingly resolved, the idea of applying a life-cycle argument to consumption-based asset pricing proved innovative. Its outstanding contribution does not necessarily lie in the fit of its parameters but instead in its identification of a more fundamental driver of the asset-market. It can be argued that Constantinides, Donaldson, and Mehra (2002) responded to Kocherlakota's call (Kocherlakota (1996), p. 87) to transcend the "current mode of patching the standard models of asset exchange with transaction costs here and risk aversion there"−quite successfully so.

Conclusion

The importance of the equity premium puzzle is that it manifests the empirical failure of the class of general equilibrium asset-pricing models. This has been of tremendous concern ever since because these and related models are still considered a cornerstone of modern macroeconomics. Changing preference-structures can actually deepen the puzzle while at the same time highlighting the role of abstractions that contribute to it. Other attempts at modifying the utility function by incorporating habit-style preferences are successful, but their ad-hoc nature lacking axiomatic foundation needs to be resolved to gain full acceptance. A particular, encouraging proposal abandons the complete-markets representative-agent approach by introducing borrowing constraints and heterogeneity in the context of an overlapping generations model. The special appeal lies with its very intuitive yet far-reaching application of the life-cycle argument to asset pricing, identifying a fundamental feature of the market for assets. Altogether, as concluded by the discoverers of the puzzle Mehra and Prescott themselves (Mehra and Prescott (2008b), p. 114), "considerable progress has been made and the equity premium is a lesser puzzle than it was twenty years ago."

References

ABEL, A. B. (1990): "Asset prices under habit formation and catching up with the Joneses," *American Economic Review, Papers and Proceedings*, 80, 38–42.

AIYAGAR, S. R., AND M. GERTLER (1991): "Asset Returns with Transaction Costs and Uninsured Individual Risk," *Journal of Monetary Economics*, 27(3), 311–331.

BREALY, R. A., S. C. MYERS, AND A. FRANKLIN (2006): *Principles of Corporate Finance*. McGraw-Hill/Irwin, New York, 8 edn.

BREEDEN, D. T. (1991): "An intertemporal asset pricing model with stochastic consumption and investment opportunities," *Journal of Financial Economics*, 7, 265–296.

CAMPBELL, J. Y., AND J. H. COCHRANE (1999): "By force of habit: A consumption-based explanation of aggregate stock market behaviour," *Journal of Political Economy*, 107, 205–251.

CONSTANTINIDES, G. M. (1990): "Habit formation: A resolution of the equity premium puzzle," *Journal of Political Economy*, 98, 519–543.

CONSTANTINIDES, G. M., J. B. DONALDSON, AND R. MEHRA (2002): "Junior can't borrow: A new perspective on the equity premium puzzle," *Quarterly Journal of Economics*, 118, 269–296.

——— (2005): "Junior is rich: bequests as consumption," *Economic Theory*, 32(1), 125–155.

DAVIS, S., AND P. WILLEN (2000): "Using financial assets to hedge labor income risks: Estimating the benefits," Working Paper.

Donaldson, J., and R. Mehra (2008): “Risk-based Explanations of the Equity Premium,” in *The Handbook of the Equity Risk Premium*, pp. 37–99. Elsevier, Amsterdam.

Epstein, L. G., and S. E. Zin (1989): “Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis,” *Journal of Political Economy*, 99, 263 – 286.

FRED (2011): “3-Month Treasury Bill: Secondary Market Rate [TB3MS],” Federal Reserve Economic Data, Federal Reserve Bank of St. Louis, Board of Governors of the Federal Reserve System.

Friend, I., and M. E. Blume (1975): “The demand for risky assets,” *American Economic Review*, 65, 900–922.

Grossman, S. J., and R. J. Shiller (1981): “The determinants of the variability of stock market prices,” *American Economic Review*, 71, 222–227.

Hansen, P. H., and R. Jagannathan (1991): “Implications of Security Market Data for Models of Dynamic Economies,” *Journal of Political Economy*, 99(2), 225–262.

Heaton, J., and D. Lucas (1996): “Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing,” *Journal of Political Economy*, 104(3), 443–487.

Homer, S. (1963): *A history of interest rates*. Rutgers University Press, New Brunswick, NJ.

Ibbotson, R. G., and R. A. Singuefield (1979): *Stocks, bonds, bills and inflation: Historical returns (1926-1978)*. Financial Analysts Research Foundation.

Kocherlakota, N. R. (1996): “The Equity Premium: It’s Still a Puzzle,” *Journal of Economic Literature*, XXXIV, 42–71.

Kreps, D., and E. Porteus (1978): “Temporal resolution of uncertainty and dynamic choice theory,” *Econometrica*, 36, 185–200.

Ljungqvist, L., and T. J. Sargent (2004): *Recursive Macroeconomic Theory*. MIT Press, Cambridge, MI, 2 edn.

Ljungqvist, L., and H. Uhlig (1999): “On consumption bunching under Campbell-Cochrane habit formation,” SFB 649 Discussion Paper, Berlin, available online at http://sfb649.wiwi.hu-berlin.de/discussionPapers de.php.

Lucas, R. E. (1978): “Asset prices in an exchange economy,” *Econometrica*, 46, 1429–1445.

Mehra, R., and E. C. Prescott (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

——— (2008a): "The Equity Premium: ABCs," in *The Handbook of the Equity Risk Premium*, pp. 1–36. Elsevier, Amsterdam.

——— (2008b): "Non-Risk-based Explanations of the Equity Premium," in *The Handbook of the Equity Risk Premium*, pp. 102–114. Elsevier, Amsterdam.

Shiller, R. J. (1989, 2011): "S&P Composite Price Index Series, Dividends Accruing to Index Series, CPI Series, Market Volatility," Chapter 26, The MIT Press, available online at `http://www.econ.yale.edu/~shiller/data.htm`.

Storesletten, K., C. I. Telmer, and A. Yaron (2004): "Consumption and Risk Sharing over the Life Cycle," *Journal of Monetary Economics*, 51(3), 609–633.

Tobin, J., and W. Dolde (1971): "Wealth, Liquidity, and Consumption," Cowles Foundation Discussion Papers 311, Cowles Foundation for Research in Economics, Yale University.

Weil, P. (1989): "The equity premium puzzle and the risk-free rate puzzle," *Journal of Monetary Economics*, 24, 401–421.

Appendix

Set of admissible average risk premia and real returns

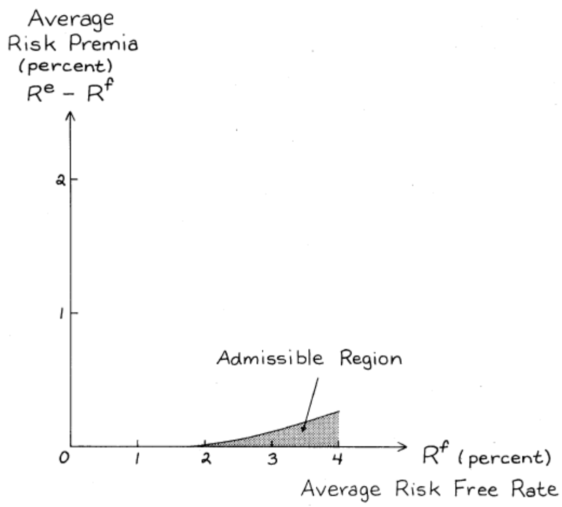

Figure 1: Set of admissible average risk premia and real returns
Source: Mehra and Prescott (1985), Fig. 4, p.155

Estimation

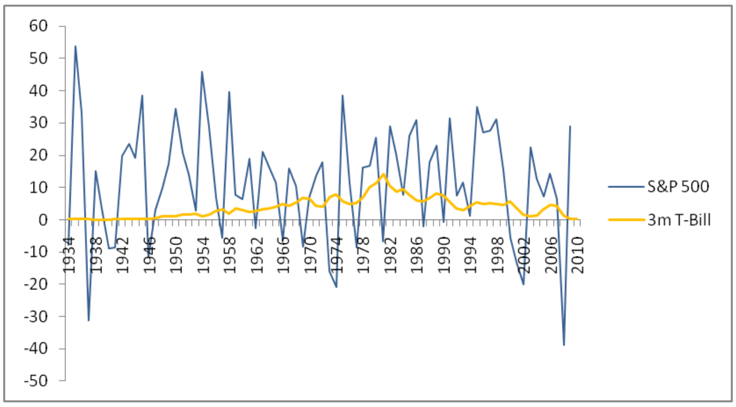

Figure 2: Mean annual total return on the S&P 500 and nominal yield on the 3-Month Treasury Bill, in percent; own calculations
Data: FRED (2011) and Shiller (1989, 2011)

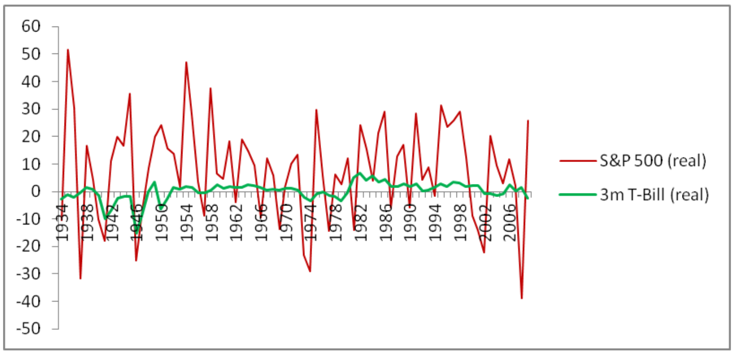

Figure 3: Mean annual real total return on the S&P 500 and real yield on the 3-Month Treasury Bill, in percent; own calculations
Data: FRED (2011) and Shiller (1989, 2011)

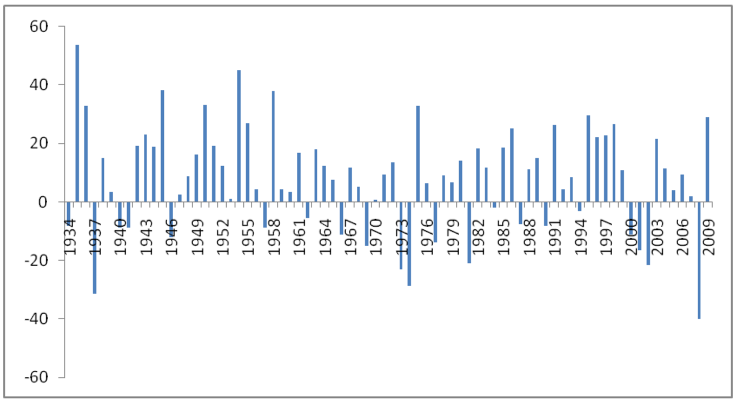

Figure 4: Annual equity premium, 1934−2010; own calculations
Data: FRED (2011) and Shiller (1989, 2011)