

The risk premium for equity: implications for resource allocation, welfare and policy*

Abstract

Analysis of the equity premium puzzle has focused on private sector capital markets. The object of this paper is to consider the welfare and policy implications of each of the broad classes of explanations of the equity premium puzzle. As would be expected, the greater the deviation from the first-best outcome implied by a given explanation of the equity premium puzzle, the more interventionist are the implied policy conclusions. Nevertheless, even explanations of the equity premium puzzle consistent with a general consumption-based asset pricing model have important welfare and policy implications.

JEL Classification: E62 H55 G12

Keywords: equity premium puzzle, resource allocation, welfare, policy

Simon Grant
Department of Economics
Rice University
and
School of Economics
Australian National University

John Quiggin
School of Economics
University of Queensland

*We thank Jan Boone, Patrick Francois, Frederic Palomino and Nancy Wallace for valuable comments and suggestions. Naturally we remain solely responsible for any remaining errors and omissions. Financial support for this project has been provided by the Australian Research Council's Large Grant A79800678. Quiggin also gratefully acknowledges income support from an ARC Federation fellowship.

The equity premium puzzle, first observed by Mehra and Prescott (1985), has given rise to a large literature. As Mehra and Prescott showed, an application of the standard consumption-based capital asset pricing model (CCAPM) with plausible parameter values implies that the risk premium associated with the market rate of return to equity, relative to the riskless bond rate, should be less than half a percentage point compared to actual values of six percentage points or more observed in the United States and other developed countries over long periods. A closely related observation is the risk-free rate puzzle (Weil 1992), that is, the fact that the real rate of returns to bonds is lower than the CCAPM prediction.

In the years immediately following the publication of Mehra and Prescott (1985), a number of proposed resolutions were discussed and rejected either because they could not be calibrated so as to fit the data or because they broke down when incorporated into a fully intertemporal version of CCAPM. More recently, the flow of papers proposing partial or complete resolutions of the puzzle has exceeded the capacity of the profession to provide critical response, but no explanation has achieved general acceptance (Mehra 2003, Mehra and Prescott 2003). Calibration exercises undertaken on models incorporating proposed explanations of the equity premium mostly yield the result that plausible parameterizations can explain part, but not all, of the observed premium. Calibrated models that have been claimed to explain the entire premium have generally not proved robust to modest changes in assumptions. Meanwhile there has been considerable interest in whether the equity premium has declined in recent years, and, if so, why. (See, for example, Jagannath, Grattan and Scherbina, 2001, and the references cited therein).

Despite the large and growing theoretical and empirical literature on the equity premium puzzle, little attention has been paid to the implications of the puzzle for resource allocation, welfare or policy. Yet the analysis of Mehra and Prescott shows that a crucial price variable, the price of systematic risk, is an order of magnitude larger than would be expected on the basis of rational optimization in efficient capital markets. As noted by Campbell and Cochrane (1999), this implies that the welfare cost of economic fluctuations is considerably greater than would be suggested on the basis of a standard CCAPM model, such as that considered by Lucas (1987).

It would be surprising if such a discrepancy between the predictions of the standard model

and reality had no other welfare or policy consequences. In fact, as will be shown below, the analysis of the equity premium is central to issues as diverse as the desirability of a ‘Tobin tax’ on international financial transactions and the evaluation of public-private partnerships under the UK government’s Private Finance Initiative.

The welfare and policy implications of the equity premium puzzle will depend on the reason for the divergence between the CCAPM prediction derived by Mehra and Prescott and the observed value of the equity premium. Three broad classes of explanations of the equity premium and risk-free rate puzzles have been offered. First, there are explanations that retain the hypothesis of rational optimization in efficient capital markets, but rely on preference structures or beliefs about the distribution of equity returns that differ from those employed by Mehra and Prescott. Second, there are explanations that rely on some form of error or misperception by investors. Third, there are explanations that invoke some form of capital market failure (relative to the CCAPM assumption of costless and complete state-contingent markets).

The object of the present paper is to consider the implications of each of these broad classes of explanations of the equity premium puzzle for resource allocation, welfare and policy. As would be expected, the greater the deviation from the first-best outcome implied by a given explanation of the equity premium puzzle, the more interventionist are the implied policy conclusions. Nevertheless, even explanations of the equity premium puzzle consistent with a general CCAPM have important welfare and policy implications.

1 Explanations of the equity premium puzzle

Long data series generally show that the rate of return to buying and holding the market portfolio of stocks is considerably greater than the rate of return to government bonds. For example, Mehra and Prescott (1985) present data showing that over the period 1889–1978, the average annual yield on the Standard and Poors 500 Index was seven per cent, while the average yield on short-term debt was less than one per cent. The difference of around six percentage points, is referred to as the equity premium.

By contrast, economic analysis based on standard models of capital markets and risk preferences, such as CCAPM, suggests that the equity premium should be small, in most cases, less than half a percentage point. The gap between the observed value of the equity premium and the value predicted by CCAPM and similar models is referred to as the "equity premium puzzle".

The equity premium is a puzzle because, in CCAPM, the equity premium is derived from risk-aversion with respect to systematic risk, that is, the correlation between returns to equity and aggregate consumption. Since aggregate consumption is quite stable (its coefficient of variation is about 3 per cent) and people appear to be only moderately risk averse, the risk premium should also be small. This argument was first developed by Mehra and Prescott, and is presented with slightly different formulations in the survey articles by Kocherlakota, Mehra and Mehra and Prescott.

Mehra and Prescott's characterization of the anomalously large equity premium as a 'puzzle' had a significant impact on the direction taken by the subsequent literature. The core research problem was presented as one of finding a satisfactory solution to the puzzle, that is, a model in which the predicted values of the equity premium and asset prices more generally matched the stylized facts noted by Mehra and Prescott. By contrast, less attention was paid to implications of the anomalous equity premium for resource allocation and for economic policy. To address the latter question, it is necessary to impose some order on the profusion of 'explanations' that have been offered as resolutions of the equity premium puzzle, by classifying them into a small number of classes. The implications of each class of explanation for resource allocation and public policy may then be considered.

A useful starting point is given by the surveys by Kocherlakota (1996), Mehra (2003) and Mehra and Prescott (2003). We exclude from consideration some proposed resolutions that are rejected by these authors on grounds that appear convincing, including the survivor bias hypotheses of Brown, Goetzmann and Ross (1995) and the 'disaster states' hypothesis of Rietz (1988) but include a number of ideas not discussed in other survivors, notably including non-optimizing models of behavior.

1.1 Individual characteristics

1.1.1 CCAPM-based explanations of the equity premium

In some sense, the most complete resolution of the equity premium paradox would be one which showed that the CCAPM is valid, even though the formulation used by Mehra and Prescott is not. The simplest explanation of the equity premium, consistent with rational optimization and efficient markets, is that the representative consumer is much more risk-averse than is normally supposed. This explanation has not found much support in view of extensive evidence supporting the view that the typical coefficient of relative risk aversion is close to 1. Nevertheless, this simple explanation has considerable heuristic value in illustrating the welfare and policy implications (if any) of the equity premium.

More plausible CCAPM-based resolutions of the equity premium puzzle include models of preferences with habit persistence (Constantinides 1990, Campbell and Cochrane 1999) or dependence on previous peak consumption (an idea first put forward by Duesenberry in the context of the debate over the Keynesian consumption function). Similarly, by relaxing the assumption of expected-utility maximization, Epstein and Zin (1990) can account for a large risk premium without necessarily requiring a correspondingly low degree of intertemporal substitution, but this again conflicts with evidence of risk-attitudes in other markets which point to much less aversion to risk.

The most important characteristic of this class of resolutions of the equity premium puzzle is that, if systematic risk in the economy arises from inherent features of the underlying technology, as in standard real business cycle models, the observed set of asset prices represents a first-best welfare outcome. Hence, any welfare or policy implications must arise as a result of the interaction between capital markets and failures in other markets. Nevertheless, intuitions about resource allocation derived, implicitly or explicitly, from the standard CCAPM may be seriously misleading.

In positive terms, explanations based on habit persistence or similar characteristics of preferences imply that short-run and long-run responses to relative price changes should be significantly

different. Indeed as Otrok, Ravikumar and Whiteman (2002) emphasize, agents whose preferences exhibit habit persistence care not only about the overall volatility of consumption but also its *temporal* distribution. Specifically, habit persistence implies much more aversion to high-frequency fluctuations than to low-frequency fluctuations in consumption. They demonstrate that a relatively small amount of high-frequency volatility in consumption can generate a large equity premium in the habit persistence model. However, they caution against such models being accepted as a complete resolution of the puzzle, since they find when compared to the actual changes in the characteristics of the frequency of volatility in US consumption over the past one-hundred years, the model's predictions for the changes in the equity premium (and risk-free rate) are essentially and significantly in the opposite direction to the observed changes.

1.1.2 Relaxation of the optimization assumptions

The standard CCAPM model represents consumers as continuously-optimizing dynamically-consistent expected-utility maximizers. These assumptions provide a tight link between attitudes to risk and attitudes to variations in intertemporal consumption. As noted above, this link may be weakened by assuming habit persistence in consumption preferences. A similar effect may be obtained by weakening the assumption of continuous optimization, for example by supposing that consumers adopt satisficing rules rather than fully optimizing.

Benartzi and Thaler (1995) propose an explanation they dub *myopic loss aversion* based on Kahneman and Tversky's (1979, 1984) two concepts of "loss aversion" (the tendency for individuals to be more sensitive to "losses" than to "gains") and "mental accounting" (the myopic heuristics people use to organize and evaluate their financial positions). Assuming investors are myopic loss averters and adjust their portfolios frequently, Benartzi and Thaler show through their simulations that the size of the equity premium is consistent with previously estimated parameters of Kahneman and Tversky's (1979) prospect theory, if investors evaluate their portfolios annually. Myopic loss aversion has also received some experimental support in Thaler et al. (1977) and more recently in Gneezy et al.(2003). In the latter paper, the 'market prices' of risky assets that arose in the various treatments were significantly higher when the feedback frequency and decision

flexibility were reduced.

Gabaix and Laibson (2001), on the other hand, argue that *delays* in adjusting consumption to take account of changes in returns to equity may lead to short-term behavior that is apparently highly risk-averse, even though the actual coefficient of risk aversion may be quite low.

In predictive terms, both of these alternative relaxations of fully rational decision-makers yield results quite similar to those obtained by Constantinides (1990). The most important distinction is that welfare analysis based on the assumption of perfect optimization may not be applicable in these models. In particular, even though the divergence from full optimization may appear small, the resulting individual welfare loss will in general be large, since observed portfolio choices yield returns well below those of the portfolio that would be chosen by a moderately risk-averse continuously-optimizing investor.

1.1.3 Error-based explanations

There has been no general acceptance of attempts to resolve the equity premium paradox while maintaining the hypothesis that individuals rationally optimize according to preferences that are, at least approximately, consistent with the assumptions of the CCAPM model. A plausible inference is that some or all individuals diverge substantially, and systematically, from the rationality assumptions of the CCAPM model. This inference leads to ‘error-based’ explanations of the equity premium puzzle.

The simplest error-based explanation of the equity premium, put forward by Glassman and Hassett (1999) is that investors, and ‘expert’ advisors, have misperceived the riskiness of equity by confusing short-term volatility with long-term risk. Glassman and Hassett argue that investors are gradually realizing that equity investments guarantee higher long-term returns with little additional risk. As they do so, the equity premium is declining. Glassman and Hassett estimate that the elimination of the equity premium would raise the price of equity by a factor of four. Since the Dow Jones index stood at 9000 when their initial estimate was made, this calculation implied the title of their bestselling book *Dow 36000*. Hence, their analysis implies that investors

can still make substantial profits by anticipating this adjustment.¹

De Long et al. (1990), drawing on the work of Shiller (1989), put forward a model in which risk over and above that due to the dividend-generating process is introduced into the economy by the mistaken, and stochastic, beliefs of noise traders, giving rise to an equity premium. A broadly similar model is presented by Cecchetti et al.(2000)

1.2 Equity characteristics explanations

A second class of explanations focuses on characteristics of equity not captured by standard capital market models, but nonetheless consistent with rational optimization by investors in smoothly functioning markets.

Principal-agent problems and the corporation The most important characteristic of equity that is not fully taken into account in standard asset pricing models is the fact that equity holders must elect managers and monitor their activities. Thus, there is a principal-agent relationship between the managers of corporations and the holders of equities.

In most cases, it is prohibitively costly for individual investors to closely monitor the activities of the corporations in which they invest. Hence, investors must rely either on external auditors and regulators or on the strong efficient market hypothesis that equity markets fully reveal even private information about asset values. The most relevant version of the latter hypothesis is that based on the idea that an uninformed investor can guarantee receipt of average market returns by investing in the market portfolio. The question of whether it is possible for an uninformed investor to identify the market portfolio is an open one. However, to explain the equity premium, it is only necessary to postulate that potential investors do not believe they can match the performance of the market.

Liquidity An alternative explanation based on characteristics of equity is put forward by writers including Bansal and Coleman (1996), Palomino (1996), Holmstrom and Tirole (1998) and Swan (2000) who focus on the demand for liquidity. Swan argues that although equity holders gain

¹ Glassman and Hassett take no account of the risk-free rate puzzle. Presumably part of the equilibrating adjustment would take the form of an increase in the riskless interest rate.

utility from frequent trade, they must be compensated, through a higher rate of return, for the transactions costs associated with equity trades.

Palomino (1996) shows that in a noise trader model with imperfect competition, the ‘thinner’ (in terms of the number of active traders) is the market for equities, the lower is its equilibrium price and hence the higher is the expected premium over the competitively priced risk-free bond. Holmstrom and Tirole (1998) develop an alternative role for liquidity in a model in which firms facing aggregate uncertainty over their liquidity needs are willing to pay a premium for Treasury bonds over private claims.

Tax distortions McGrattan and Prescott (2001) argue that the observed equity premium in the United States since 1945 may be explained by changes in the tax treatment of interest and dividend income. As Mehra (2003) notes, there are some difficulties in the calibration used in this analysis and the existence of a substantial equity premium before 1945 is left unexplained. Also as noted by Mehra, the premium for equity obtained in this way is not a risk premium, which is problematic in view of the relatively good performance of models like CAPM that yield high estimates of the risk premium. Nevertheless, the general observation that frictions arising from taxes and other distortions need to be taken into account in explaining the equity premium is an important one.

1.3 Market failure explanations

Two broad classes of market failure have been considered as explanations of the equity premium. First, problems of adverse selection and moral hazard may result in the absence of markets in which individuals can insure themselves against systematic risk in labor income and noncorporate profits. Second, transactions costs or liquidity constraints may prevent individuals from smoothing consumption over time.

Kocherlakota rejects explanations of the equity premium based on the absence of insurance markets, arguing that individuals could overcome the consequences of uninsurable risk by borrowing and lending to smooth consumption over time. It seems plausible, however, that a model incorporating both the absence of contemporaneous insurance markets and constraints on in-

tertemporal consumption smoothing would be consistent with the observed behavior of asset prices.

1.3.1 Insurance market failures

Mankiw (1986) and Weil (1992) argue that the presence of this uninsurable background risk may reduce willingness to hold additional systematic risk in the form of equity and thereby account for an equity premium. Mankiw stresses the point that, *ex post*, systematic risk is concentrated on the relatively small group in the population who incur unemployment or business failure. If individuals dislike negatively skewed returns (that is, in expected-utility terms, if they display prudence, so that $u''' > 0$), then the equilibrium risk premium will be higher when *ex post* risk is concentrated on a subset of the population.

In Weil's model, individuals who are making their portfolio decisions are *ex ante* identical in terms of the distribution of risks that they face. But the risk associated with their labor income is assumed to be idiosyncratic and uninsurable. Equity pays out a dividend that is assumed to be independently distributed to an individual's labor income. If they lack the ability to pool their labor income, individuals are more adverse to holding equity. Weil shows that the relative price of equity to bonds is smaller when individuals cannot pool their labor income risk. Moreover, the absolute price of bonds is higher, implying a lower return risk-free rate. That is, undiversifiable or uninsurable background risk both increases the (relative) equity premium and reduces the risk-free rate.²

1.3.2 Constraints on intertemporal consumption smoothing

Constantinides, Donaldson and Mehra (2002) present an overlapping generations model, with three generations, in which the young, who have limited financial resources of their own, and who would otherwise issue debt and purchase equity, are constrained from borrowing. The old wish to sell their entire portfolio to finance their consumption, which only leaves the middle-aged to hold both equity and the positive net supply of bonds. At given asset prices, this depresses demand

² More precisely, Weil (1982) proves that both $E[(1+r+\rho)/(1+r)]$ and $E[r]$ increase when individuals are unable to pool their labor income. As Grant and Quiggin (2000) observe, this is not the same as proving that $E[\rho]$ must increase. They provide an example with constant relative risk averse preferences where $E[\rho]$ decreases.

for equity and raises the demand for bonds. Hence, both a higher bond price (implying a lower risk-free rate) and a larger equity premium are required for equilibrium in this economy compared to unconstrained economy in which young do not face borrowing constraints.

2 Implications for resource allocation

2.1 The economic cost of risk

The most direct and obvious implication of a large risk premium for equity is that the economic cost of the systematic risk in returns to equity is also large. This is also the only implication of the equity premium that has received extensive discussion. Responding to the argument of Lucas (1987) that the standard CCAPM model yields very small estimates of the welfare cost of economic fluctuations, particularly when compared to the benefits of economic growth, Campbell and Cochrane observe that, in a model that generates values of the equity premium consistent with observed asset prices, the welfare costs of systematic risk are comparable in magnitude to the welfare benefits of economic growth. Although this result is model-specific, the general point that a large risk premium for equity implies a large welfare cost in any model where asset prices reflect preferences (Atkeson and Phelan 1994).

For illustrative purposes, we will assume, conservatively, that the real bond rate is 2 per cent and that the expected return to the market portfolio is 8 per cent, so that the equity premium is 6 percentage points. It follows that any stationary stream of returns to equity with the same risk characteristics as the market portfolio has a market value one-quarter of that of a similar stream of returns with no systematic risk. Equivalently, in flow terms, the welfare cost of systematic risk in equity returns, evaluated at prevailing market prices, is equal to three-quarters of the value of corporate profits, or around 15 per cent of GDP, assuming a profit share of 20 per cent. Presumably, if risk were reduced, the equity premium would also decline, implying that the counterfactual of a complete elimination of systematic risk in equity returns would raise welfare by less than 15 per cent. Nevertheless, the impact is large.

Assuming that equity is completely characterized by its state-contingent distribution of returns, this result follows directly from the existence of the equity premium, whether or not the premium

reflects market failure or incomplete optimization. However, if the premium arises from other characteristics of equity, such as illiquidity or principal-agent problems, returns to equity cannot be used to determine the economic cost of risk.

The implications for the economic costs of risk may also be considered in state-contingent terms. In the simplest possible model, there are two states of nature, normal and recession. We think of a bond as a state-contingent security yielding a (normalized) unit return in both states of nature and a unit of equity as a security that also yields a unit return in the normal state but yields a lower return k in the recession nature. We calibrate the model in order to match the stylized facts regarding the rates of return to bonds (0.02) and equity (0.08), and the standard deviation of return to equity (0.2).

In this framework, the state-contingent price of income in the recession state (expressed in terms of state-independent income claims) must satisfy

$$q - p = (1 - k)p_r,$$

where q is the price of a bond, p is the price of equity and p_r is the marginal value of recession-state income.

Let π denote the probability of the recession state. The rates of return to debt and equity are given by

$$\begin{aligned} r &= 1/q \text{ and} \\ \rho &= ((1 - \pi) + \pi k) / p, \end{aligned}$$

respectively, and the standard deviation of risky equity returns is given by

$$\sigma = \sqrt{(1 - k)^2[\pi(1 - \pi)]/p^2}.$$

If we assume $\pi = 0.1$, we can solve these four equations for the unknowns p , q , k and p_r . With the stated parameters, we obtain $q = 50$, $p = 6.8$, $k = -3.5$ and $p_r = 9.6$. Hence p_n , the marginal value of normal-state income is equal to $q - p_r = 40.4$. The probability-adjusted relative price of recession-state income in terms of normal-state income, is therefore given by

$$\frac{p_r}{p_n} \times \frac{(1 - \pi)}{\pi} = 2.1$$

That is, at the margin, an increase in income of one unit in the recession state, would be exactly offset by a reduction in income of 2.1 units in the normal state. Equivalently, relative to the contributions to expected income, marginal income units in a recession year are worth around twice as much as marginal income units in a normal year.

2.2 Individual portfolio decisions

There is a large, and influential, popular literature dealing with the implications of the equity premium for individual portfolio decisions. The predominant tone of this literature has been ‘bullish’, presenting the argument that the historically high rates of return to equity are evidence of excessive or misplaced risk-aversion, and that long-term investors can earn high returns, with little or no risk, by switching to stocks from bonds and other assets. A notable contribution is that of Glassman and Hassett (1999), who argued that the correction of erroneous beliefs would lead to a fourfold increase in the value of stocks, raising the Dow Jones index from 9000 to 36000.

The contrary view has been presented most prominently by Shiller (2000). Shiller begins by noting that the central thesis of Glassman and Hassett was almost universally accepted by the late 1990s. Shiller quotes a survey, conducted in 1999, in which 96 per cent of respondents agreed with the statement ‘The stock market is the best investment for long-term holders who can just buy and hold through the ups and downs of the market’, with 76 per cent indicating strong agreement and only 2 per cent disagreeing.

Shiller does not offer an explicit explanation for the equity premium. However, in view of his earlier work on the excess volatility of stock prices, it is natural to link his analysis with a ‘noise trader’ model like that of De Long et al. (1990). In these models, it is possible for well-informed rational investors to make positive expected profits by exploiting the erroneous beliefs of noise traders. Faced with a bubble such as that of 1996–2000, Shiller advises rational investors (presumably those lacking the deep pockets required for short-selling) to remain on the sidelines, and hold a substantial portion of wealth in assets other than equity.

Both the Shiller and Glassman–Hassett models are based on market inefficiencies arising from persistent error on the part of some market participants. It follows that rewards will accrue to

investors who can identify and exploit the prevalent patterns of error. By contrast, explanations of the equity premium that are consistent with the efficient market hypothesis yield no guidance with respect to portfolio choices. By definition, any available information about market behavior is already reflected in the prices generated by an efficient market.

The same appears to be true of market failure and transaction-costs models, in which the prices generated by equity markets are typically assumed to eliminate any opportunities for arbitrage, when information constraints and transactions costs are taken into account.

2.3 ‘Short-termism’

In the 1980s and early 1990s, there was vigorous debate over claims that the dominant role of stock markets in investment decision-making in English-speaking countries generates a bias towards investments with a ‘short-term’ bias. By contrast, it was claimed Japanese and European financial systems, in which banks played a more active role, resulted in a greater supply of ‘patient capital’. Interest in the ‘short-termism’ debate receded in response to the strong growth in output and investment in the United States in the late 1990s, and the poor performance of the Japanese economy over the same period, but many of the issues remained unresolved.

An influential case in support of the short-termism thesis was put forward by Porter (1990, 1992), while opposing views were argued by Jensen (1986) and Marsh (1990). Most theoretical discussion of the problem focused on informational asymmetries between investors and corporate managers (Naranayan 1985, Shleifer and Vishny 1990). However, as Miles (1993) observes, the problem of short-term bias is related to (though not identical to) the equity premium puzzle.

Supposing that the risk characteristics of firms are stable and that cash flows follow a diffusion process with the conditional variance of flows proportional to their maturity, the CAPM model implies that any enterprise i can be valued by discounting its cash flow using a risk-adjusted discount rate $r_i = r + \beta\rho$ where β measures the covariability of returns with the market portfolio. Hence, the larger is ρ , the larger is r_i and the shorter is the corresponding value of the ‘payback period’ commonly used in project evaluation (if mean cash flows are constant, the payback period is $\frac{1}{r_i}$).

Particular explanations of the equity premium puzzle may also imply the existence of systematic deviations from CAPM valuations for a given value of the equity premium ρ . For example, if investors misperceive the riskiness of longer-term investments (say those with returns more than five years in the future) this will be reflected in an equity premium that is disproportionately large for firms undertaking such investments. Miles (1993) finds evidence of such a pattern.

Models based on capital market failure may also imply systematic deviations from CAPM valuations. For example, if individuals are confident about short-term macroeconomic conditions, but uncertain about conditions in the medium and longer term, their inability to insure against the consequences of recession would be reflected in a short-term bias, relative to the predictions of a CAPM model.

2.4 Trends in the equity premium

A number of writers have suggested that the equity premium has declined, or even disappeared, in recent years (Jagannath, Grattan and Scherbina 2001). This claim is obviously consistent with the view of Glassman and Hassett (1999) that the equity premium arose from errors which are being corrected over time. A number of other proposed explanations of the equity premium puzzle yield a similar prediction, given appropriate auxiliary hypotheses. For example, if transactions costs in securities markets are declining as a result of deregulation or technological progress, models based on liquidity costs or market frictions imply that the equity premium should decline. As noted by McGrattan and Prescott (2001), a declining trend in taxes should produce a temporary equity premium which will ultimately disappear. Similarly, if transactions costs in markets for insurance and intertemporal consumption smoothing are falling, market failure explanations are consistent with a declining equity premium. It seems clear that transactions costs in securities markets have declined, as have taxes on capital income. It is less clear that the costs of personal insurance and consumption-smoothing transactions have fallen.

By contrast, explanations of the equity premium based on individual characteristics such as loss aversion or habit persistence give no reason for predicting a trend in the equity premium. In general, it would be expected that the distribution of such characteristics in the population would

be broadly stable over time.

Finally, ‘noise trader’ models predict a cyclical equity premium which declines in periods when noise traders are excessively optimistic, and increases when optimism is replaced by pessimism. On this view, the decline of the equity premium since 1980 would be seen as the product of a temporary bubble in stock prices.

3 Welfare and policy implications

3.1 Recessions and fiscal policy

Perhaps the most important and robust welfare and policy implication of resolutions of the equity premium puzzle is that the welfare cost of recessions is large, even if the long-term growth rate of aggregate consumption is unaffected. This result may be contrasted with the analysis of Lucas (1987) who shows that, in an expected-utility model, the welfare cost of variance in aggregated consumption is trivial.

This is true whether the explanation of the equity premium puzzle is based on market failures, on mistakes or on high effective risk-aversion. The crucial observation is that the risk premium associated with the market portfolio is determined by the relative prices of claims on income in recession and non-recession states. The larger the risk premium, the larger the marginal value of recession-state income and the greater the welfare cost of recessions. The calculations presented above provide one approach to quantifying these impacts.

This result has obvious implications for the calculation of optimal ‘sacrifice ratios’, that is, marginal-tradeoffs between output loss and reductions in inflation. There are also implications for policies such as microeconomic reform, which promise long-term gain (higher productivity) in return for short-term pain (adjustment costs). There has been a long-standing debate over whether such policies should be pursued during periods of recession. The analysis above suggests that, other things being equal, it is preferable not to incur transaction costs during recessions, but to defer them to periods of normal activity. The analysis also provides a way of quantifying some of the relevant trade-offs.

3.2 Social security

Although the welfare and policy implications of the equity premium puzzle have received little systematic attention, the divergence between expected rates of return to equity and to debt has influenced a number of contributors to the debate over reform of Social Security in the United States. This debate has been prompted in part by the observation that, under the existing pay-as-you-go arrangements, the retirement of the baby boom generation will lead to a requirement to supplement the Social Security fund from general revenue. Two suggestions based on the observed equity premium have attracted particular attention. The Clinton administration proposed a diversification of the investments of the Social Security trust fund. A committee appointed by the Bush administration has advocated the establishment of individualized accounts which would allow each person paying into social security to decide on the investment portfolio for part of their contributions. Their own future social security entitlements would then be based partly on the performance of their individual accounts. Notice that both proposals involve social security contributions being invested in equity as well as (government-issued) debt but the former scheme *pools* the returns while in the latter scheme individuals face the risk of their own investment strategy.

Grant and Quiggin (2002) show that the market failure explanations of the equity premium puzzle proposed by Mankiw (1986) and Weil (1992) imply that welfare would be improved by a Clinton-style diversification of the investment policy of the Social Security fund to include equity as well as debt. But in their model no such benefit would arise from individualized accounts, since the portfolio choices made by individuals for their social security accounts would be offset by the reallocation of their privately-held assets. This reallocation would leave the equilibrium contingent consumption of individuals unchanged.

It should be emphasized that the benefit that Grant and Quiggin identify is *not* the result of higher expected returns from equity. Indeed, the government's purchase of equity arising from the diversification of the social security trust fund would increase equity prices and reduce expected returns. Rather it arises from the fact that the government has the power to levy taxes in the future to meet any obligation it has to make any specific defined payment in the future. Hence when the government invests in equity today, in states of the world tomorrow where equity returns are high,

taxes can be lower than they otherwise would have been to meet the social security entitlements while in bad states where the equity returns are poor, taxes need to be raised. In an economy where individuals face undiversifiable risk associated with adverse selection problems, and where taxes are proportional or progressive, this means that from the perspective of an individual today, the spread of the distribution of his or her after-tax income has been increased for a boom event in the future but reduced for a recession event. In conjunction with the usual assumption that agents exhibit decreasing absolute risk aversion, Grant and Quiggin show that ex ante welfare is increased since, in utility terms, the loss from the given increase in risk at higher incomes is lower than the gain from reducing risk at lower levels of income.

A welfare improvement will also arise under the hypothesis that the equity premium arises from a mistaken perception of the riskiness of equity, provided that this misperception does not extend to the indirect holdings of equity arising from the changed investment policy of the Social Security fund. Assuming that individuals disregard risk associated with possible variations in tax rates associated with fluctuations in equity (correctly, since the risk is small relative to total income) but overestimate the risk associated with personal holdings of equity, an increase in the expected returns to the investments of the Social Security fund will be treated as equivalent to a reduction in the expected future taxes needed to meet unfunded liabilities.

Similar arguments apply with respect to explanations of the equity premium based on undesirable characteristics of equity as an asset. Supposing, for example, that the equity premium arises from the superior liquidity of debt, as in Swan (2000), and that the liquidity or otherwise of the Social Security fund investment portfolio is not a matter of concern, a shift in the holdings of the Social Security fund from debt to equity is equivalent to an increase in the supply of liquidity, and therefore generates a welfare benefit.

By contrast, under CCAPM-based explanations of the equity premium puzzle, and assuming that investors are well-informed, changes in the investment policy of the Social Security fund will be negated by offsetting changes in individual asset demands (Geanakoplos, Mitchell and Zeldes, 1998).

3.3 Public investment and privatization

The issue of risk premiums and their implications for public and private investment was debated during the 1960s and 1970s. Hirshleifer (1964) argued that the use of the risk-free rate to discount the returns of public projects would lead to the displacement of superior private projects. On the other hand, Arrow and Lind (1970), argued that, because the public sector had a superior capacity to spread risk, the value of a risky project was greater under public ownership than under private ownership and reaffirmed the view that benefit–cost analysis for public projects should be conducted using the riskless rate of discount.

The formal argument of Arrow and Lind (1970) showed that there should be no risk premium for public investment in a small project with returns uncorrelated with aggregate income. More significantly, in their policy discussion, Arrow and Lind implicitly assumed the central results of Mehra and Prescott (1985), namely that, under standard assumptions, the risk premium associated with perfect diversification of all risks in the economy would be close to zero but that the risk premium actually required for private equity investment was large. By contrast, in the absence of explicit estimates of the optimal risk premium, critics of the Arrow–Lind proposition, such as Hirshleifer (1989) began from the presumption that capital markets operated efficiently to spread risk. This led to the conclusion that the market-determined price of systematic risk was socially optimal and should be used in the evaluation of public projects.

This issue is of considerable policy relevance in the light of widespread use of a ‘public sector comparator’ in the evaluation of public-private partnerships proposed under the Private Finance Initiative in the United Kingdom and similar schemes. The use of the CAPM model to evaluate the risk borne by the public sector comparator has been the subject of considerable controversy (see for example, Mayston 1999, Officer 1999 and Quiggin 1999).

Reconsideration of the debate in the light of the literature on the equity premium paradox clarifies a number of points. Under the standard CCAPM assumptions, including perfect capital markets, the optimal rate of discount for public projects will be close to the riskless rate (as claimed by Arrow and Lind) and equal to the discount rate for private projects with similar risk characteristics (as claimed by Hirshleifer). The central disagreement between the two sides rests

on the fact that while Arrow and Lind implicitly assumed that an analysis similar to that of Mehra and Prescott could be applied to yield the optimal rate of discount for public projects, Hirshleifer and other critics of Arrow and Lind implicitly assumed that the observed cost of capital for private firms was consistent with CCAPM.

Thus, the Arrow–Lind argument relies implicitly on the existence of a market failure, but Arrow and Lind (1970) do not specify how the market failure arises. If risk-spreading through the tax system is to correct the market failure, it is important to focus on problems arising from adverse selection rather than moral hazard. Since taxation is not voluntary, tax policies can overcome adverse selection problems. However, moral hazard problems based on unobservable effort cannot be overcome through tax policy.

Grant and Quiggin (2003a) show that, if adverse selection problems prevent insurance against systematic risk, as in Mankiw (1986), the optimal rate of return for public projects will be less than the market rate for projects with similar risk characteristics. Grant and Quiggin consider the case where the public sector has available a menu of projects arranged in decreasing order of attractiveness relative to the market portfolio. They show that, for the marginal project, the rate of return will be greater than the riskless bond rate but less than the private-sector rate of return.

The Arrow–Lind proposition is also valid if the equity premium arises from characteristics of equity not incorporated in the standard CCAPM model. This point is most evident in the model of Swan (2000) where the equity premium arises from transactions costs associated with trade in imperfectly liquid equity. In this model, asset prices are determined optimally and the real bond rate is equal to the social opportunity cost of capital. By virtue of its superior ability to issue a liquid security the government enjoys a cost advantage relative to issuers of private equity. Hence, the appropriate rate of discount for public projects is the bond rate.

The case where the equity premium arises from errors raises some difficulties in welfare analysis. If, as Shiller (1989) argues, financial markets display excess volatility, then returns to holdings of equity are riskier than are the associated streams of corporate profits. Shiller’s insight has been formalized in the ‘noise trader’ model of De Long et al. (1990). In this model, risk over and above that due to the dividend-generating process is introduced into the economy by the distorted

and stochastic beliefs of noise traders. De Long et al. observe that this excess risk implies an increase in the equity premium relative to the case when all investors have rational expectations, and they show that, although both noise traders and sophisticated investors are made better off in *ex ante* terms (given their beliefs) by the availability of trade, this apparent welfare improvement arises at the expense of those holding equity when trade is introduced, such as entrepreneurs making initial public offerings or governments privatizing publicly owned assets. With regard to a privatization program, if the equity premium arises from the mistaken beliefs of noise traders, then as Grant and Quiggin (2003b) show the privatization may reduce public sector net worth. Moreover, evaluated in terms of the correct beliefs of sophisticated investors, there is a reduction in social welfare associated with privatization that must be balanced against any improvements in operating efficiency.

3.4 Tobin taxes

The proposal of Tobin for a tax on international financial transactions, designed to reduce the volatility of exchange rates and facilitate macroeconomic management has attracted considerable attention (see for example, ul Haq, Kaul and Grunberg, 1996). Stiglitz (1992) has presented similar arguments for taxes on domestic asset market transactions. Given the possibility of arbitrage, it seems likely that any practical proposal would require both international and domestic transactions to be taxed at the same rate.

As Palley (1999) observes, a noise trader model implies that taxes should reduce volatility and increase welfare (on average, given correct beliefs). Empirical tests are difficult since removal of taxes also increases the equilibrium price of equity directly for any positive level of turnover, so would want to use revenue to reduce taxes on profits. By contrast, Swan (2000) argues that, in a liquidity model, taxes on transactions will reduce welfare. Since the equity premium in models of this kind is generated by transactions costs, this claim appears plausible.

4 Concluding comments

The CCAPM is a powerful and intuitively appealing model of asset price determination, based on assumptions which seem to correspond reasonable well to reality. No-one would suppose that

these assumptions were perfectly accurate. Consumer-investors are not perfectly rational, utility is not perfectly additively separable over time, there are costs associated with a range of financial transactions, some risks are uninsurable and so on. Applications of the model have proceeded on the basis of the standard methodological presumption that none of these violations of the assumptions are individually large enough to render the model inapplicable or to induce large biases in its predictions, and that taken collectively, a large number of independent violations of the model assumptions will approximately cancel out.

As far as relative asset prices are concerned, the standard presumption appears to be accurate. The CCAPM is widely employed and seems to work reasonably well in practice. By contrast, the equity premium puzzle shows that, using standard parameters and a standard model setup, the CCAPM prediction of the premium associated with systematic risk is out by an order of magnitude. The development of the subsequent literature indicates that ‘resolutions’ of the CCAPM based on changes in a single parameter or modelling assumption have not resolved the puzzle to the satisfaction of the economics profession.

It seems reasonable to infer that the violation of the CCAPM predictions reflects two problems. First, most violations of the assumptions underlying the CCAPM tend to raise rather than lower the equity premium. Second, there are positive interaction effects, such that the combined impact of multiple violations of the assumptions is greater than the individual impact of any one violation. In retrospect, this is not surprising. Equity markets are social contrivances for spreading risk. The assumptions of costless transactions, complete state-contingent markets and additive expected-utility preferences are particularly favorable to risk-spreading. A model based on these assumptions is therefore likely to yield lower estimates of the market price of risk-bearing than more complex and realistic models.

Given that the resolution of the equity premium puzzle is likely to involve interactions between a number of deviations from the assumptions of the CCAPM model, assessment of the welfare and policy implications of the puzzle raises some significant difficulties. The most robust implications of the equity premium puzzle are those that flow directly from the high price of systematic risk and are therefore independent of the resolution of the puzzle. For example, there can be little doubt

that the existence of a large equity premium strengthens the case for macroeconomic stabilization policy. Issues relating to the cost of capital for the public sector, including social security policy and the analysis of privatization, are more complex. To the extent that the equity premium is generated by factors that do not affect the government, such as adverse selection problems and costs associated with financial transactions, it seems reasonable to infer that the cost of capital will be close to the rate predicted by CCAPM, and therefore close to the real bond rate. On the other hand, if the equity premium is consistent with a modified CCAPM, or arises from problems that also apply in the public sector, the public cost of capital will be close to the corporate cost.

References

- Arrow, K. and Lind, R. (1970), “Uncertainty and the evaluation of public investment decisions”, *American Economic Review* 60(2), 364-78.
- Bansal, R. and Coleman, J. (1996), ‘A monetary explanation of the equity premium, term premium, and risk-free rate puzzles’, *Journal of Political Economy*, 104(6), 1135–71.
- Brown, S., Goetzmann, W. and Ross, S. (1995), ‘Survival’, *Journal of Finance*, 50, 853 – 87.
- Benartzi, S. and Thaler, R. (1995), “Myopic loss aversion and the equity premium puzzle”, *Quarterly Journal of Economics*, 110, 73-92.
- Campbell, J. and Cochrane, J. (1995), ‘By force of habit: a consumption-based explanation of aggregate stock market behavior’, NBER Working Paper 4995.
- Cecchetti, S., Lam, P.- S., and Mark, N. (2000), “Asset pricing with distorted beliefs: Are equity returns too good to be true?”, *American Economic Review* 90(Sept.), 787-805.
- Constantinides, G. M. (1990), “Habit formation: a resolution of the equity premium puzzle”, *Journal of Political Economy* 98(3), 519-43.
- Constantinides, G. M. and Duffie, D. (1996), “Asset pricing with heterogeneous consumers”, *Journal of Political Economy* 104(2), 219-40.
- Constantinides, G. M., Donaldson, J. B., and Mehra, R. (2002), “Junior can’t borrow: a new perspective on the equity premium puzzle”, *Quarterly Journal of Economics* 117(1), 269-296.
- De Long, J. B., Shleifer, A., Summers, L., and Waldmann, R. J. (1990), “Noise trader risk in financial markets”, *Journal of Political Economy* 98(Aug.), 703-738.
- Epstein, L. G. and Zin, S. E. (1989), “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework”, *Econometrica* 57, 937-69.
- Epstein, L. G. and Zin, S. E. (1990), “First-order’ risk aversion and the equity premium puzzle”, *Journal of Monetary Economics* 26(3), 387-407.
- Gabaix X. and Laibson, D. (2001), “The 6D bias and the equity premium puzzle”, MIT Working Paper 02-01.

- Geanakoplos, J., Mitchell, O. S. and Zeldes, S. P. (1998), "Would a privatized social security system really pay a higher rate of return?" unpublished mimeo.
- Glassman, J. and Hassett, K. (1999), *Dow 36000*, Times Books, New York.
- Gneezy, U., Kapteyn, A. and Potters, J. (2003), "Evaluation periods and asset prices in a market experiment", *Journal of Finance* 58(2), 821-37.
- Gollier, C. (2001), *The Economics of Risk and Time*, MIT Press, Cambridge MA.
- Grant, S. and Quiggin, J. (2000), "The interaction between the equity premium and the risk-free rate", *Economic Letters* 69(1), 71-79.
- Grant, S. and Quiggin, J. (2002), "The risk premium for equity: implications for Clinton's proposed diversification of the social security fund", *American Economic Review* 92(5), 1104-1115.
- Grant, S. and Quiggin, J. (2003a), "Public investment and the risk premium for equity", *Economica* 70(277), 1-18.
- Grant, S. and Quiggin, J. (2003b), "Noise trader risk and the and the welfare effects of privatization", Rice University Working Paper 2003-03.
- Greenspan, A. (1999), Testimony before US Senate Budget Committee, January.
- Grossman, S.J. and Shiller, R.J. (1982), "Consumption Correlatedness and Risk Measurement in Economies with Non-Traded Assets and Heterogeneous Information", *Journal of Financial Economics* 10(2), 195-210
- Grossman, S. J., Melino, A. and Shiller, R.J. (1987) "Estimating the Continuous-Time Consumption-Based Asset-Pricing Model", *Journal of Business and Economic Statistics* 5(3), 315-27.
- Hirshleifer, J. (1964), "Efficient Allocation of Capital in an Uncertain World", *American Economic Review*, 54(3), 77-85.
- Hirshleifer, J. (1989), *Time, uncertainty and information*, Basil Blackwell, New York.
- Holmstrom, B. and Tirole, J. (1998), "Private and public supply of liquidity", *Journal of Political Economy* 106, 1-40.
- Jagannathan, R., McGrattan, E. R. and Scherbina, A. (2001), Federal Reserve Bank of Minneapolis *Quarterly Review* forthcoming.
- Jensen, M. (1986), "The takeover controversy: analysis and evidence", *Midland Corporate Finance Journal* 4(1), 6-32.
- Kahneman, D. and Tversky, A. (1979), "Prospect theory: an analysis of decision under risk", *Econometrica* 47, 263-291.
- Kahneman, D. and Tversky, A. (1984), "Choices, values and frames", *American Psychologist* 39, 341-50.
- Kocherlakota, N. R. (1996), "The Equity Premium: It's Still a Puzzle", *Journal of Economic Literature* 34(1), 42-71.
- Lucas, R. (1987) *Models of Business Cycles*, Basil Blackwell, New York.
- Mankiw, N. G. (1986), "The equity premium and the concentration of aggregate shocks", *Journal of Financial Economics* 17, 211-19.
- Marsh, P. (1990), Short-termism on trial, International Fund Managers Association, London.
- Mayston, D. (1999), 'The Private Finance Initiative in the NHS: an unhealthy development in New Public Management?', *Financial Accountability and Management*, 15(3/4), 249-74.

- McGrattan, E. and Prescott, E. (2001), ‘ Taxes, regulations, and asset prices’, Federal Reserve Bank of Minneapolis, Working Paper 610.
- Mehra, R. (2003), ‘The equity premium: why is it a puzzle?’, NBER Working Paper, 9512.
- Mehra, R. and Prescott, E. C. (1985), “The equity premium: a puzzle”, *Journal of Monetary Economics* 15(2), 145-61.
- Mehra, R. and Prescott, E. (2003), ‘The equity premium in retrospect’, NBER Working Paper, 9525.
- Miles, D. (1993), “Testing for short termism in the UK stock market”, *Economic Journal* 103(421), 1379-96.
- Naranayan (1985), “Observability and the payback criterion”, *Journal of Business* 58(July), 309-24.
- Orosel, G. O. (1996), “Informational efficiency and welfare in the stock market”, *European Economic Review* 40, 1379-1411.
- Otrok, C., Ravikumar, B. and Whiteman, C. H. (2002), “Habit formation: a resolution of the equity premium puzzle?” *Journal of Monetary Economics* 49(6), 1261-1288.
- Officer, R. (1999), ‘Privatisation of public assets’, In *Privatisation: Efficiency or Fallacy? Two Perspectives*, *Growth*, 50 1–22.
- Palley, T. (1999), “Speculation and Tobin taxes: Why sand in the wheels can increase economic efficiency”, *Zeitschrift Fur Nationalokonomie* 69(2), 113-126.
- Palomino, F. (1996), “Noise trading in small markets,” *Journal of Finance* 51, 1537-1550.
- Porter, M. E. (1990), *The competitive advantage of nations*, Free Press, N.Y.
- Porter, M. E. (1992), “Capital disadvantage: America’s failing capital investment system”, *Harvard Business Review* (Sep/Oct), 65-82.
- Quiggin, J. (1999), ‘The future of public ownership in Australia: Privatisation or renationalisation?’, In *Privatisation: Efficiency or fallacy? Two perspectives*, *Growth*, 50, 23–50.
- Quiggin, J. and Chambers, R. G. (1998), “A state-contingent production approach to principal-agent problems with an application to point-source pollution control”, *Journal of Public Economics* 70, 441-72.
- Rietz, T. (1988), ‘The equity risk premium: A solution’, *Journal of Monetary Economics*, 22(1), 117–31.
- Shiller, R. J. (1989), *Market volatility*, MIT Press, Cambridge, Mass. and London.
- Shiller, R. J. and Grossman, S. J. (1981), “The determinants of the variability of stock market prices”, *American Economic Review* 71, 222-27.
- Shiller, R. J., Melino, A. and Grossman, S. J. (1987), “Estimating the continuous-time consumption-based asset-pricing model”, *Journal of Business and Economic Statistics* 5, 315-27.
- Shliefer, A. and Vishny, R. (1990), “Equilibrium short horizons of investors and firms”, *American Economic Review* 80(2), 148-53.
- Stiglitz, J. (1989), “Using tax policy to curb speculative short-term trading”, *Journal of Financial Services Research* 3, 101-115.
- Stiglitz, J. and Weiss, A. (1981), “Credit rationing in markets with imperfect information”, *American Economic Review* 71, 383-410.

Swan, P. (2000), "Does 'illiquidity' rather than 'risk aversion' explain the equity premium puzzle? The value of endogenous market trading", No. 18 December, School of Business, University of Sydney.

ul Haq, M., Kaul, I. and Grunberg, I., eds. (1996) *The Tobin tax: Coping with financial volatility*, Oxford University Press, New York and Oxford.

Weil, P. (1992), "Equilibrium Asset Prices with Undiversifiable Labor Income Risk." *Journal of Economic Dynamics and Control* 16 (3-4), 769-90.