Journal of Economic Dynamics & Control 37 (2013) 1736–1754

Contents lists available at SciVerse ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Returns-to-scale and the equity premium puzzle

Geoffrey Dunbar [a,b,*]

[a] Vancouver School of Economics, University of British Columbia, 997-1873 East Mall, Vancouver BC, Canada V6T 1Z1
[b] Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada V5A 1S6

ARTICLE INFO

Article history:
Received 9 June 2009
Received in revised form
16 November 2012
Accepted 23 January 2013
Available online 30 April 2013

JEL classification:
E23
E43
E44
G11
G12

Keywords:
Equity premium
Risk-free rate
Returns-to-scale

ABSTRACT

A model of heterogenous firms facing idiosyncratic risk is proposed which generates an equity premium of 6 per cent and a risk-free rate of 1.5 per cent even if aggregate returns are risk-free. The premium in this model reflects diminishing returns-to-scale and the fact that equity shares are equal claims to firm output. In the bond market, the risk-free rate reflects trade in assets at marginal rates of return with a linear technology and thus the equity premium in excess returns reflects a comparison of average returns with marginal returns. In the model, credit constraints lower the equity premium and, absent such constraints, the equity premium would roughly double. Since the model may be interpreted as a model of entrepreneurship, this paper also presents estimates from a structural model of entrepreneurship using data from the Survey of Consumer Finances and also finds only a modest level of risk aversion is sufficient to replicate entrepreneurial returns.

1. Introduction

The equity premium puzzle, identified by Mehra and Prescott (1985), refers to the sizeable excess return of equities over the returns to a risk-free asset – an apparent anomaly that has been the inspiration of much quantitative and theoretical research. Attention has focused primarily on the difference in the risk characteristics of equities and treasury bills (the usual proxy for a risk-free investment) and the role that investors' attitudes towards risk play in explaining the premium.[1] Certainly, risk attributes are one difference between both assets and so the increased variance in equity returns should require higher returns to attract investment if investors are risk averse. Yet, standard asset pricing models have difficulty matching the excess return of equities given plausible degrees of risk aversion. Perhaps fortunately, there is second, often overlooked, difference between both assets that may help to explain the equity premium. Realized equity returns are measured empirically as average returns to equity (for example, the denominator is typically a total equity valuation) and so factors which lead to differences in average returns may shed light on the equity premium puzzle. One such factor, and the focus of this paper, is returns-to-scale.

* Correspondence address: Vancouver School of Economics, University of British Columbia, 997-1873 East Mall, Vancouver BC, Canada V6T 1Z1. Tel.: +1 604 822 2876; fax: +1 604 822 5915.
E-mail addresses: geoffrey.dunbar@ubc.ca, dunbar.geoffrey@gmail.com

[1] See Mehra (2008), Mehra and Prescott (2003) and Kocherlakota (1996), for instance, Jagannathan et al. (2001) present some evidence that the equity premium fell substantially in the latter part of the 20th century. This evidence does not, however, explain the occurrence of the puzzle in the preceeding period.

0165-1889/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jedc.2013.04.007

Certainly there is some reason to believe that the equity premium puzzle may reflect something more than investors' attitude towards risk. Moskowitz and Vissing-Jørgensen (2002) examine the private equity premium, the equity premium for privately held firms, in the U.S. and find that it is almost identical to that of publicly held firms. Yet privately held firms typically represent a large part of an entrepreneur's wealth and are subject to greater investment uncertainty than a diversified basket of equities. Theoretically, one might expect that entrepreneurs would earn a higher return to compensate for the additional risk. The evidence suggests that they do not.

In this paper, a fixed measure of heterogeneous firms are endowed with a diminishing returns-to-scale production technology, own capital and are subject to idiosyncratic production risk but no aggregate production risk. The absence of aggregate risk is by design, both for tractability and expository purposes. The absence of aggregate risk implies that any 'premium' obtained in the model cannot be explained as compensation to a diversified investor for risk. The specific question addressed is whether heterogeneity and diminishing returns-to-scale production can help explain the level of the equity premium for low levels of investor risk aversion. The answer appears to be yes.

If equity returns face diminishing returns-to-scale, then the average return is greater than the marginal return. Diminishing returns also imply that the size of firm equity matters for the average return – equity returns are decreasing in firm size which is consistent with the empirical evidence (e.g. Fama and French, 1992 and Barber and Lyon, 1997). The risk associated with total returns matters in two ways for average returns: it affects the optimal size of the firm equity and it affects the price of risk-free instruments. Thus, the combination of idiosyncratic risk and diminishing returns-to-scale may help to resolve the equity premium puzzle without requiring highly risk-averse investors. Indeed, in this paper, investors share a constant relative risk aversion parameter of 1.5 and there is nevertheless an equity premium of 6.7 per cent and a risk-free rate of 1.5 per cent. This premium and risk-free rate are consistent with post-war US data as reported by Shiller (1989).

Shareholders are assumed to own the capital stock of the firm through their equity holdings and share the total return of a firm *equally*. Since equity shares are identical claims to the capital stock and output of the firm, they each receive the average return to a share. Indeed, one key to the results in this paper is the assumption that firms *own* capital rather than renting it from households (as in Akdeniz and Dechert, 2007 for example) but that households own firms. The ownership of capital by firms implies that capital re-allocation between firms may be constrained by the credit market. Thus, internal rates of return for capital vary between firms depending on each firm's ability to borrow or lend in the credit market and due to managerial preferences over dividend streams. These differences in internal rates of return can imply implausibly large leverage ratios and so, for most of the analysis, I assume that firms are constrained to have non-negative shareholder book-value—they cannot borrow more than the present value of their assets. This restriction appears to match empirical facts. It may be also supported by the observation that shareholders face no liability claims and so cannot support non-negative valuations in a rational framework.

Since firms in the model face occasionally binding borrowing constraints, the capital allocation across firms is inefficient. One natural question to pose is how reducing the borrowing constraints on firms would change equity returns. In the model, one may interpret joint management of firms (i.e. mergers) as alleviating the borrowing constraint since under joint management firms are not constrained in their (re-)allocation of capital.[2] Using the same parameter values as in the numerical example mentioned above yields an equity premium of 12.8 per cent and a risk-free rate of 4.17 per cent when all firms are jointly managed. Thus, in the model, mergers are likely to increase equity returns which mirrors the conclusions of the empirical literature, e.g. Martynova and Renneboog (2008). Matching this fact is typically difficult in neoclassical models because mergers diversify dividend income risk and so should result in lower risk premia. Here, mergers increase the marginal returns to capital but increase the average return by a larger degree owing to the diminishing returns to scale technology. This implies that credit market constraints lower equity premia that result from diminishing returns to scale and firm heterogeneity.[3]

Although most neoclassical production models of the equity premium embed diminishing returns to capital, they do so in the context of a constant-returns-to-scale technology and competitive factor markets in which a factor unit is paid its marginal product. Indeed, two (common) modeling assumptions impose the restriction that the marginal product of all units of capital is constant. (1) Firms have a linear production technology in capital and all capital is paid the same market clearing rate, e.g. Gomes et al. (2003), and (2) production is constant-returns-to-scale and factors are paid their marginal product, e.g. Cochrane (1991). Both (1) and (2) are assumptions. Much research in macroeconomics does not assume or imply either (1) or (2), for instance: labor-search, managerial quality/span of control, timing of productivity shocks within periods and incomplete markets. If assumptions (1) or (2) do not hold, equity returns may depend on the level of firm capital. In this paper, I am silent on the underlying mechanism which yields diminishing returns to scale and instead evaluate the extent to which diminishing returns may explain the equity premium.

Another literature, including Meh and Quadrini (2006), Covas (2006), and Cagetti and De Nardi (2006) explicitly model the entrepreneurial sector of the economy as having decreasing-returns-to-scale to the single factor, capital, that entrepreneurs use to generate output. It is common in this literature to construct the private-equity risk premia as

[2] Joint management does not imply that capital stocks are combined but rather that capital is reallocated efficiently between firms. In this light, it might be more appropriate to view firms as plants (or service delivery platforms). So joint-management between Walmart and Sam's Club, for example, implies that capital is allocated to maximize the joint return by equalizing the marginal returns at both.

[3] It also implies that credit market constraints help to explain the low risk-free rate puzzle identified by Weil (1989).