Risk, return and self-employment *

Harvey Rosen[†]
Princeton University
and NBER

Paul Willen[‡]
Graduate School of Business
University of Chicago

July 2, 2002

Abstract

Do people avoid self-employment because it is too risky? To address this question, we begin by measuring the level and variance of household income for both wage-earners and the self-employed using individual level panel data. We then construct a theoretical model of the choice between the two occupational modes. We use the theoretical model, in conjunction with our estimates of the level and variance of income, to assess the extent to which occupational choice is driven by differences in tolerance toward risk. The model explicitly takes into account the possible role of portfolio investment in attenuating the uncertainty associated with labor income. Our main finding is that the increase in mean consumption that rewards the increased variance of self-employment is much too large to be rationalized by conventional measures of risk aversion. This result is robust to a number of alternative specifications of the model, including one that explicitly takes survivorship bias into account. We conclude that willingness to accept risk is not a dominant factor in the decision to become self-employed.

1 Introduction

In an address to Princeton University's graduating class of 2002, entrepreneur Meg Whitman, CEO of eBay, told the students, "You must bravely engage opportunity

*We are grateful to Mehmet Tosun for assistance in putting together the dataset, to seminar participants at the University of Chicago for useful comments and Princeton's Center for Economic Policy Studies for support.

†Phone: (609)258-4022. E-mail: hsr@princeton.edu

‡Phone: (312)946-0829. E-mail: paul.willen@gsb.uchicago.edu

and accept risk...Take risks, undertake new endeavors, even though you may not succeed." Whitman's exhortation reflects the view, found in both popular and academic discourse, that entrepreneurship and risk-bearing are inextricably linked. As Brockhaus (1980, p. 513) notes, "Entrepreneurial literature since the writings of Mill has included risk bearing as a major distinguishing characteristic... of an entrepreneur."[1] This is due to the fact that "...in becoming an entrepreneur an individual risks financial well-being, career opportunities, family relations, and psychic well-being" (Brockhaus, 1980, p. 510). The notion that entrepreneurs are relatively tolerant of risk is a common assumption in theoretical models of occupational choice (Knight (1921), Kihlstrom and Laffont (1979), Kanbur (1979) and Blanchflower and Oswald (1998)). The title of a recent paper characterized as a "truism" the idea that low risk aversion encourages people to become entrepreneurs (Cramer, *et al.* (2001)).

What factual basis is there for the notion that entrepreneurship is risky? It is certainly well-documented that entrepreneurial ventures have high rates of failure. For example, in the Panel Study of Income Dynamics data, about 35 percent of individuals who become self-employed exit within one year (Quadrini, 1999). Furthermore, the incomes of entrepreneurs exhibit considerably more volatility than those of wage-earners with comparable characteristics. (See, for example, Borjas and Bronars (1989), Hamilton (1995) and Heaton and Lucas (2000).) But the fact that both risk and return are higher in self-employment does not by itself tell us that differences in tolerance for risk drive the choice between self-employment and wage-earning. One need think only of the literature on the equity risk-premium puzzle to see why. Both risks and returns are higher for equity than for bonds, but when economic theory is used to analyze the differentials, it turns out that risk aversion cannot explain investors' decisions with respect to holdings of equity. Similarly, one needs a theoretical framework to determine whether tolerance for risk explains the risk-return tradeoff implicit in people's choices between self-employment and wage-earning.

In this paper we examine the issue of the riskiness of entrepreneurship through the lens of modern work on life-cycle consumption and portfolio choice. We begin in Section 2 by characterizing the income-generating processes for both self-employed and wage-earning individuals. The estimates are based on data drawn from the Panel Study of Income Dynamics for the years 1968 to 1993. Using this approach, we confirm earlier findings that self-employed households have higher mean income and higher income volatility than wage and salary households.

In Section 3 we develop a model for assessing whether the empirical results from Section 2 are consistent with the notion that attitudes toward risk are the major

[1] Indeed, Cramer *et al.* (2001) trace the notion back even earlier, to the eighteenth century.

determinant of the choice between wage-earning and self-employment. A key insight of the model is that each occupation is associated with a whole set of combinations of the level and variance of lifetime consumption, which we call the $C - V$ feasible set. In the spirit of the traditional literature on the decision to become an entrepreneur, individuals who are observationally equivalent choose among the bundles in the same feasible set on the basis of their preferences toward risk. We develop a method for constructing each occupation's $C - V$ feasible set using information about the individual income process for that occupation and asset returns. In Section 4, we use the model, in conjunction with the estimates from Section 2, to show that the benefit of higher consumption generated by self-employment outweighs the cost of higher variance for all but pathologically risk-averse consumers. That is, the model predicts that essentially everyone should choose self-employment. This result is robust to a number of alternative specifications of the model, including one that explicitly takes survivorship bias into account.

Thus, a model based on the key assumption in most of the entrepreneurial literature–that differences in risk aversion drive the choice between wage-earning and self-employment–leads to a result that is totally at odds with the data (three-quarters of our sample never even *try* self-employment). We treat this inconsistency between the model and the data as a rejection of the crucial assumption that all (observationally equivalent) individuals face the same opportunity set and that risk-aversion alone determines occupational choice. Indeed, we argue that risk aversion is unlikely to be an important factor, let along the dominant factor, in the occupational choice decision.[2] Thus, the historical tendency to identify entrepreneurship and tolerance for risk may be misplaced. Section 5 concludes with a summary and suggestions for future research.

2 Income generating processes for wage-earners and the self-employed

2.1 Preliminaries

The starting point for investigating whether individuals with relatively high tolerance for risk choose self-employment is measuring the mean and variance of the returns

[2]Newman (1995) comes to a conclusion that is similar in spirit using an approach entirely different from ours. He investigates entrepreneurship in the context of a theoretical model of optimal incomplete insurance contracts and finds that risk-based explanations for entrepreneurship are implausible.

in each occupation.[3] This, in turn, requires estimates of the respective occupations' income generating processes. To estimate these processes, we need individual panel data that include, *inter alia*, information on income and employment status. The Panel Study of Income Dynamics (PSID) is suitable for these purposes; we use the waves from 1968 to 1993. Our unit of observation is the household. We focus on households in which the heads were between the ages of 25 and 64 inclusive.

An important issue is how to define income. If everyone were single and we knew what part of self-employment income derived from capital and what part from labor, then the answer would be clear–use the individual's labor income. However, neither of these conditions is true. Instead, in our canonical model, we focus on the *household's* total income, including capital income. We focus on family income for four reasons. First, decisions about occupations are made in a family context–it is the volatility of *family* consumption and income that matters to the individual. Second, any distinction between labor and capital income in the context of an integrated family business is inherently arbitrary. Third, the opportunity to invest in a high-return project may be one of attractions of self-employment. All of this said, it is of some interest to see whether the results change when we focus on the head's earnings rather than the family's income. We show below that our substantive results are not very sensitive to such a change in the definition of income. All figures are converted to 1996 levels using the personal consumption expenditure deflator.

The designation of employment status is not entirely straightforward. In practice, as already noted, it is not uncommon for an individual to be self-employed for a few years and then return to wage-earning. Indeed, as implied by the summary statistics in Table 1, 36 percent of the men in our sample who tried self-employment were self-employed only one year. The comparable figure for women was 57 percent. Our approach is to classify an individual as self-employed only if he or she was self-employed a substantial amount of time. Clearly, some arbitrariness is involved in defining "substantial." In our basic model, we classify individuals as being self-employed only if they were self-employed five years or more during the period we observe them. We experimented with several other cutoff points, and found that our qualitative results were not affected.[4]

[3]In contrast, Moskowitz and Vissing-Jorgensen (2001) compare returns to investment in entrepreurial ventures with returns to publicly-traded equity. They find that the returns are similar.

[4]A related issue is that a very few individuals report being both wage-earners and self-employed during a given year. We classified these individuals as being self-employed for that year; construing them as being wage-earners had no impact on the results. Note that self-employed individuals need not have any capital investment in their businesses. Indeed, according to the tabulations by Hurst and Lusardi (2001), only fifty percent of those describing themselves as self-employed actually own

2.2 Income Generating Equation

Our income generating model is a second order autoregression in first differences.[5] Because standard portfolio theory suggests that the riskiness of either occupation depends on the correlation of its return with financial asset returns, we augment the autoregressive terms with the real rate of return on equity and its lag:

$$\tilde{y}_t - y_{t-1} = \tilde{\eta}_t + \alpha_0 + \chi_1 (y_{t-1} - y_{t-2}) + \chi_2 (y_{t-2} - y_{t-3}) + \phi_0 \tilde{r}_t + \phi_1 r_{t-1} \quad (1)$$

where y_t is real income, η_t is an innovation, $\tilde{r}_t$ is the real rate of return on the value-weighted New York Stock Exchange composite index and r_{t-1} is its lag, and α_0, χ_1, χ_2, ϕ_0, and ϕ_1 are parameters. There is some evidence that income generating processes differ by education, so we estimate separate equations for those individuals with no more than high school and those with some college or more education. [6] The regression equations also include a cubic in age. The equations are estimated with individual effects, so that there is no need to control for unchanging individual characteristics such as race.

The income generation parameters and associated calculations are reported in Table 2. Within each education category, the first row shows the figures for wage and salary individuals and the second row shows comparable figures for the self-employed. Consider first the column labeled "mean/(median)," which shows the mean value of income in thousands of 1996 dollars and its median in parentheses. These are raw means and medians for individuals between the ages of 35 and 40. Looking at a specific age group is a simple way to account for the fact that the age distributions of the self-employed and wage-earners differ. Within each education category, the self-employed have higher mean household incomes than wage-earners. The medians for the self-employed are also higher within each category, but the differences are smaller. This is a common result (see, e.g., Hamilton (2000)), and is due to the substantial skewness of the incomes of the self-employed.

The next column has two figures in each cell. The first is the standard deviation of the innovation; the second is the ratio of the standard deviation to the mean value of income. Consistent with previous work (Hamilton (2000)), the income stream

businesses and of business owners, 38 percent have business equity of less than $5000. Moskowitz and Vissing-Jorgensen (2001) and Heaton and Lucas (2000) focus on business owners, in contrast with our focus on the self-employed. Neither group is a subset of the other.

[5] Previous work indicates that there is a unit root in a levels specification, so that first-differencing is appropriate. See, for example, MaCurdy (1982).

[6] It might also be of interest to stratify the sample by occupation in order to examine, for example, differences in the income generating processes of self-employed and wage-earning physicians. However, sample sizes were not sufficiently large to permit such analysis.

associated with self-employment is more variable than that associated with wage-earning, both in absolute terms and relative to income.[7]

Next consider the coefficients χ_1 and χ_2, which multiply the first and second lagged differences in income, respectively. Both lags are significant for all education classes and employment classes. In general, the greater the sum of χ_1 and χ_2 in absolute value, the less persistent the process. For households in which the head has no more than a high school education, the income generating process is somewhat more persistent for the self-employed; and vice versa for those with some college or more.

The columns headed β_{nyse} and $\beta_{nyse_{lag}}$ indicate the relationship between earnings and the current and lagged rates of return on financial assets, respectively. The results of a joint test of the significance of β_{nyse} and $\beta_{nyse_{lag}}$ are included in the column headed F_{nyse}. The first number in each cell is the F-statistic and the second number is the associated p-value. For wage earners, there is some tendency toward a negative correlation between earnings and asset returns. For both self-employed and wage-earning individuals with no more than high school, the β's are jointly statistically significant, and their sum is negative. Thus, individuals in this group can reduce the variance of income by investing in financial assets, although the effect is stronger for the self-employed. For those with some college or more, the point estimates suggest that the relationship between financial asset returns and earnings is positive for wage-earners, but marginally significant. For the self-employed, the point estimate is negative, but not statistically significant.

2.2.1 Alternative specifications

We have shown that, within both education classes, self-employment income is higher than the income of salaried workers, but also more variable. We next examine whether these findings are robust to a number of alternative approaches to measuring income and self-employment status.

To begin, we consider an alternative definition of income. Specifically, rather than looking at household income, we analyze the earnings of the head. As noted above, we think that there are good reasons to focus on the volatility of all of the household's income streams. But because much of the literature focuses on earnings, it is useful

[7] It has been suggested that the standard deviation of labor income in the PSID is overestimated because of measurement error (Deaton (1991)). To the extent this is true, it makes self-employment more attractive relative to wage-earning and hence reinforces our key finding below that issues relating to risk do not seem to be the driving force in the choice between the two occupational modes.

to examine the outcome when we use them to fit the income generating function. For the sake of brevity, we do not report the entire set of results. Rather, in the bottom half of Table 3 we show, by occupation and education, the mean and median earnings, the standard deviation of earnings, and its coefficient of variation. The patterns with respect to mean household income and its variance that are present in the top half of the Table are here as well; of course, the magnitudes are greater in the top half of the table because family income is generally greater than individual earnings. Thus, our stylized facts are robust to a reasonable change in how income is defined.

Another issue concerns the fact that there is some arbitrariness to our procedure for classifying people as self-employed, which is based on whether or not they were self-employed five or more years. We experimented with a number of different cutoffs ranging from 2 years to 10 years. The substantive results from Table 2 are not affected by such changes in the criterion for classifying individuals as self-employed. The results for the 10-year cutoff are particularly noteworthy, and are reported in Table 3. Presumably, the individuals who survive 10 or more years in self-employment are relatively successful at it. Nevertheless, the variance of their life-cycle incomes is still high. This suggests that our finding of a relatively high variance in the canonical specification, Table 2, is not due to the fact that a lot of the individuals classified as self-employed are failures who would have high-variance incomes in any case. Rather, self-employment income appears to be inherently volatile.

But there is a potential problem with any approach based on such cutoffs. In effect, such approaches select on "winners" – individuals who are able to stick it out as self-employed for several years. To deal with this issue, we implement the following algorithm: Determine whether or not an individual was self-employed before the age of 35. Then, estimate the income generating equation using data only from years *after* the individual was 35, but classify the individual on the basis of his or her self-employment status *before* the age of 35. Thus, inclusion in the sample used to estimate the self-employment income generating function does not depend on whether one is self-employed during any of the years used in the estimation. This approach creates no survivorship bias, that is, sample selection in favor of successful self-employed people. Roughly speaking, this is like an instrumental variables procedure in which the instrument is a dichotomous variable for whether or not the individual was self-employed prior to the time period used to estimate the regression. The results from this "pseudo instrumental variables" approach are reported in Table 4. The results are quite similar to those in our basic model, suggesting that they are not driven by survivorship bias.

The approaches taken so far allow for the fact that individuals can move between self-employment and wage-earning. It will turn out to be useful also to consider a tack based on the opposite assumption, that the choice is in some sense irrevocable. Specifically, consider the following classification algorithm: once an individual experiences a spell of self-employment, no matter how brief, *all* the subsequent years are used to estimate the self-employment income generating equation, whether or not the individual was actually self-employed. In effect, the assumption is that any exposure to self-employment leads to an irrevocable change in the income generating process. The results from this "irrevocable" approach are included in Table 4. They indicate that it does not materially affect our substantive findings.

For easy reference, Table 3 contains a summary of all the results we have discussed so far, as well as a few other variations. The basic message is that the qualitative results are quite consistent across specifications.

3 Theoretical Framework

In the previous section we computed the moments of the income processes for self-employment (defined in various ways) and wage-earning. The key stylized fact that emerges is that self-employment is associated with both a higher mean and variance of income. Now, standard theoretical considerations suggest that this is entirely predictable–people need a higher expected return in order to compensate for taking on additional risk. But the key question in this context is whether the amount of compensation is reasonable given conventional assumptions about risk aversion.

In this section we develop a theoretical model of the choice between wage earning and self-employment that allows us to investigate this issue. Our model building strategy is aimed at providing a simple and transparent framework for interpreting the empirical findings from Section 2. The idea is to see how far a model that focuses on risk can take us in explaining the facts documented in the previous section. Our assumptions with respect to the form of the utility function and the distribution of returns are designed to lead to specifications that can easily be solved analytically. A key simplifying assumption is that the self-employment decision is irrevocable. As noted in the previous section, this is clearly counterfactual. However, as also noted in the previous section, even short-lived experiences with self-employment can be viewed as having irrevocable consequences for future earnings streams. Viewed in this way, a theoretical model that assumes irrevocability can in fact cast light on actual self-employment decisions.

In our model, each occupational mode is entirely characterized by an uncertain stream of future income, so that the selection of the preferred income stream is, in effect, the selection of an occupation. The individual selects the income stream that maximizes the expected value of an intertemporal utility function that depends on period by period consumption. We show that, under certain conditions, the choices available to the individual can be characterized very simply. Following much recent work, the model explicitly takes into account the possible role of portfolio investment in attenuating the uncertainty associated with labor income (see Davis and Willen (2000a, 2000b)). To fix ideas, we begin with a two period model, and then turn to the more general case.

3.1 Two-period model

Consider an individual who lives for two periods ($t = 0, 1$). Initially, she has no financial assets. But she can choose from two lifetime income profiles, one associated with wage-earning and one with self-employment, and which are indexed by i. Each lifetime income profile delivers current income y_0^i with certainty and stochastic period one income $\tilde{y}_1^i$. In addition, she has access to two financial assets: *asset 0* is a riskless bond with certain gross return (i.e., one plus the rate of return) R_0; *asset 1* is a risky security with uncertain gross return $\tilde{R}_1$. We assume that labor income innovations, $\tilde{\eta}_1^i = \tilde{y}_1^i - \mathrm{E}(\tilde{y}_1^i)$, and risky asset returns $\tilde{R}_1$ are jointly normally distributed. Our individual invests ω_0 dollars in the riskless asset and ω_1 dollars in the risky asset. Let c_0 and $\tilde{c}_1$ be consumption in periods zero and one, respectively and call the pair $(c_0, \tilde{c}_1)$ a lifetime consumption profile.

3.1.1 Characterizing the budget constraint

Assume our individual chooses profile i. Asset holdings, lifetime income profiles and consumption are related by the following period-by-period budget constraint:

$$c_0 = y_0^i - \omega_0 - \omega_1 \tag{2}$$

$$\tilde{c}_1 = \tilde{y}_1^i + \omega_0 R_0 + \omega_1 \tilde{R}_1 \tag{3}$$

Equations (2) and (3) imply the following intertemporal budget constraint in expectations:

$$c_0 + \frac{1}{R_0}\mathrm{E}(\tilde{c}_1) = y_0^i + \frac{1}{R_0}\mathrm{E}(\tilde{y}_1^i) + \frac{1}{R_0}\Big(\mathrm{E}(\tilde{R}_1) - R_0\Big)\omega_1 \tag{4}$$

To simplify, call $C = c_0 + \frac{1}{R_0}\mathrm{E}(\tilde{c}_1)$ lifetime consumption, $Y^i = y_0^i + \frac{1}{R_0}\mathrm{E}(\tilde{y}_1^i)$ lifetime income. Further, let $ER = (1/R_0)\Big(\mathrm{E}(\tilde{R}_1) - R_0\Big)$ and call $ER\omega_1$ lifetime excess returns. Using the new notation, we can write equation (4) much more compactly:

$$C = Y^i + ER\omega_1 \tag{5}$$

Equation (5) tells us that lifetime consumption for a consumer must equal lifetime labor income for the selected income profile i plus lifetime excess returns.

3.1.2 Characterizing the $C - V$ indifference Curves

Turning now to preferences, let the lifetime utility function over c_0 and $\tilde{c}_1$ be time-separable and assume that each period's utility function has the exponential form $\frac{1}{A}\exp(-Ac)$ where $A > 0$. This functional form implies constant absolute risk aversion, which is given by A. As a convenience, assume that the subjective discount rate equals the riskless rate. Under these conditions, we can characterize the present discounted value of utility as a function of lifetime consumption and the variance of period one consumption:

$$U(C, V) = -\frac{1}{a_0 A}\exp\left\{-a_0 A\left(C - \frac{A}{2}V\right)\right\}, \tag{6}$$

where A measures absolute risk aversion, a_0 is an annuitization factor and $V = \mathrm{var}(\tilde{c}_1)$. [8] While a considerable simplification, the assumption of absolute risk aversion allows for closed-form, intuitive, and transparent solutions. Further, as shown below, our results are so dramatic that it is hard to imagine that a more general functional form would make much of a difference.

Equation (6) implicitly defines the individual's indifference curves between lifetime consumption and variance. In a graph with lifetime consumption (C) on the y-axis and the variance of second period consumption (V) on the x-axis, utility increases as we move up and to the left. To generate an indifference curve, choose a real number u. Any consumption profile that satisfies

$$C - \frac{A}{2}V = u \tag{7}$$

delivers the same level of utility, so that equation (7) defines an indifference curve. Two features of these indifference curves merit attention. First, implicit differentiation of equation (7) implies that indifference curves in $C - V$ space have slope $\frac{A}{2}$. Intuitively, the slope of the indifference curve reflects the consumer's willingness to trade off the level of consumption (C) and the variance of consumption (V). In effect, this is a measure of the individual's risk aversion. The fact that the slope is constant reflects the underlying assumption of constant absolute risk aversion. From a geometric point of view, this means that the indifference curves are parallel to each other. Further, the higher the risk aversion, the steeper the indifference curves. Intuitively, the greater the value of A, the more consumption the individual requires to compensate her for assuming another unit of risk.

[8] Davis and Willen (2000b) provide a derivation of equation (6).

All of this is illustrated in the top panel of Figure 1. There we consider the case of an individual who can obtain lifetime consumption of \$80,000 with no risk. The figure reflects the fact that if the individual has absolute risk aversion of 0.025 (corresponding to local relative risk aversion of 1.0), she would demand \$1200 in lifetime consumption in order to take on additional consumption variance of \$100 million (standard deviation of \$10 thousand). On the other hand, if her value of A is 0.250 (corresponding to local relative risk aversion of 10), she requires \$12,500 in compensation for the higher variance.

As already suggested, with constant absolute risk aversion, the indifference curves for a given investor are parallel, as displayed in the lower panel of Figure 1. With other preferences, absolute risk aversion and thus the curvature of the indifference curves depends on both the level of lifetime consumption and the consumption variance.[9]

3.1.3 Characterizing the $C-V$ feasible sets

Just as we can characterize the individual's preferences in consumption-variance ($C-V$) space, we can characterize her feasible combinations as well. To find the feasible points in $C-V$ space corresponding to any income profile i, we take advantage of the intertemporal budget constraint (equation (4)). The budget constraint tells us that the expected lifetime consumption and variance associated with each income profile depend on ω_1, the amount of saving in the risky asset. For example, we know that the $C-V$ feasible set must include the point $(Y^i, \text{var}(\tilde{y}_1^i))$, because one can attain it simply by investing nothing in the risky asset (i.e. setting $\omega_1 = 0$). As we vary ω_1 from zero in either direction, we generate other points in $C-V$ space.

A particularly important value of ω_1 is the negative of the beta of labor income with the risky asset, β_y^i, defined as $\text{cov}(\tilde{y}_1^i, \tilde{R})/\text{var}(\tilde{R})$. When $\omega_1 = -\beta_y^i$, the associated point in $C-V$ space, $(C_{mv}^i, V_{mv}^i) = (Y^i - \beta_y^i ER, \text{var}(\tilde{y}_1^i - \beta_y^i \tilde{R}))$, has lower consumption variance than any other feasible point. (This is because β_y^i is the coefficient in the regression of $\tilde{y}_1^i$ on $\tilde{R}$; by construction it minimizes the variance of the residual.) We therefore call this point the *minimum variance point*. Note that it makes intuitive sense that if β_y^i is zero, the minimum variance point is the same as the point at which there is no investment in the risky asset: if labor income and the risky asset are uncorrelated, then investment in the risky asset cannot reduce lifetime consumption variance.

Remarkably, once the expected return and variance of the risky financial asset are given, all information relevant to the investor about the feasible set is contained in

[9]Specifically, for the common isoelastic specification (constant relative risk aversion), the slope of the indifference curves rises with the variance of consumption and decreases with wealth.

the minimum variance point. In other words, conditional on the characteristics of the risky financial asset, if two different labor income profiles have the same minimum variance point, then they generate exactly the same feasible set. To show this, take the second-period budget constraint (equation (3)) and add and subtract $\beta_y \tilde{R}$ to the right hand side to obtain:

$$V = \text{var}(\tilde{c}_1) = \text{var}(\tilde{y}_1^i + \omega_1 \tilde{R}) = \text{var}\left(\underbrace{\tilde{y}_1^i - \beta_y^i \tilde{R}}_{\text{residual}} + \left(\omega_1 + \beta_y^i\right) \tilde{R} \right) \tag{8}$$

$\tilde{y}_1^i - \beta_y^i \tilde{R}$ is the residual from a regression of income on the risky asset and is thus uncorrelated with returns on the risky asset. Therefore

$$V = \underbrace{\text{var}(\tilde{y}_1^i - \beta_y^i \tilde{R})}_{=V_{mv}^i} + \left(\omega_1 + \beta_y^i\right)^2 \text{var}(\tilde{R}) \tag{9}$$

Thus,

$$\omega_1 + \beta_y^i = \frac{\left(V - V_{mv}^i\right)^{1/2}}{\text{std}(\tilde{R})} \tag{10}$$

We can similarly decompose C:

$$C = \underbrace{Y^i - \beta_y^i ER}_{=C_{mv}^i} + \left(\omega_1 + \beta_y^i\right) ER \tag{11}$$

Equation (11) implies that:

$$\omega_1 + \beta_y^i = \frac{C - C_{mv}^i}{ER} \tag{12}$$

Any feasible point (C, V) must satisfy both equations (10) and (12). Hence, by setting the right hand sides of the two equations equal to each other, we know that the function

$$C = \left(V - V_{mv}^i\right)^{1/2} \underbrace{\frac{ER}{\text{std}(\tilde{R})}}_{\text{Sharpe ratio}} + C_{mv}^i \tag{13}$$

characterizes the $C - V$ feasible set. As promised, conditional on the characteristics of the risky financial asset (as summarized by the Sharpe ratio), the feasible set that corresponds to a given labor income profile i is fully characterized by the minimum variance point.

Geometrically, Equation (13) is a parabola that opens to the right with vertex at the minimum variance point. At the minimum variance point, an individual invests $-\beta_y^i$ in the risky asset. Note that as the Sharpe ratio increases, the greater the amount of expected lifetime consumption the individual can obtain for taking on more risk.

Intuitively, the greater the reward for taking on a unit of risk, the more that lifetime consumption increases for each unit increase in variance.

Incidentally, this line or reasoning provides a further justification for our focus in Section 2 on total family income (including capital income), as opposed to earned income only, when estimating the income generating process. If capital income consists of the return on investment in traded equity or is perfectly correlated with traded equity, then whether we include or ignore capital income is irrelevant. To see this, suppose that in the initial period, the family borrows ϕ dollars which it then invests in the risky asset. Family income in the initial period is then:

$$y_0^i = y_{L,0}^i$$

In period one, income is the sum of labor income plus the return on the risky asset minus repayment of the loan

$$\tilde{y}_1^i = \tilde{y}_{L,1}^i + \phi(\tilde{R} - R_0)$$

Let $\beta_{y,L}^i$ and Y_L^i represent the beta of labor income alone and the value of lifetime labor income alone (i.e. when $\phi = 0$), respectively. It is easy to see that:

$$Y^i = Y_L^i + \phi ER \tag{14}$$

$$\beta_y^i = \beta_{y,L}^i + \phi \tag{15}$$

Substituting equation (14) and (15) into the definitions of C_{mv}^i and V_{mv}^i yields

$$C_{mv}^i = Y_L^i + \beta_{y,L}^i ER \tag{16}$$

$$V_{mv}^i = \mathrm{var}(\tilde{y}_L^i - \beta_{y,L}^i \tilde{R}) \tag{17}$$

Equations (16) and (17) illustrate that we get the same measures of C_{mv}^i and V_{mv}^i whether or not we include capital income.

3.1.4 Portfolio choice

For the moment, assume that the individual has chosen to be a wage-earner. To maximize utility, she chooses the point in the feasible set for wage-earning that is tangent to the highest indifference curve. The top panel of Figure 2 illustrates such an equilibrium for specific numerical values of the parameters of the model. In this example and all that follow, we assume that $R_0 = 1$, $ER = .08$ and $\mathrm{std}(\tilde{R}) = .15$. Further, we assume $A = .075$ (which corresponds to a local relative risk aversion coefficient of approximately 3) here and in the other examples unless specifically noted. The characteristics of the occupations we use in our various examples are listed in Table 5. The characteristics of the income stream associated with wage earning are listed in the row marked "WS" in the table.

In the figure, O_i is the optimal point and MV_i is the minimum variance point. As usual, (C_i, V_i) maximizes utility if and only if the slope of the indifference curve (equation (7)) equals the slope of the $C - V$ feasible set (equation (13)). Formally:

$$\frac{A}{2} = \frac{1}{2}\left(V_i - V^i_{mv}\right)^{-1/2} \frac{ER}{\text{std}(\tilde{R})} \tag{18}$$

Equation (18) implies an optimal level of V conditional on occupation i:

$$V_i = V^i_{mv} + \frac{ER^2}{(A)^2 \,\text{var}(\tilde{R})} \tag{19}$$

If we substitute the optimal level of V into equation (13), we obtain the optimal level of C conditional on occupation i:

$$C_i = C^i_{mv} + \frac{ER^2}{A \,\text{var}(\tilde{R})} \tag{20}$$

From equations (19) and (20) we see that as the return to the risky asset increases *ceteris paribus*, the individual takes advantage of it to choose a lifetime profile with higher variance and higher consumption. Similarly, if the variance of the risky asset increases or the individual becomes more risk averse, both the mean and variance of the lifetime profile fall.

3.1.5 Occupational choice

Our individual chooses between two occupations $i = SE$ (self-employed) and $i = WS$ (wage and salary earner). She chooses wage-earning over self-employment if and only if

$$U\left(C_{WS}, V_{WS}\right) \geq U\left(C_{SE}, V_{SE}\right) \tag{21}$$

By equation (6), equation (21) holds if and only if

$$C_{WS} - (A/2)V_{WS} \geq C_{SE} - (A/2)V_{SE} \tag{22}$$

If we substitute the optimal solutions conditional on occupation choice i (equations (19) and (20)), we can re-write equation (22):

$$C^{WS}_{mv} + \frac{ER^2}{A \,\text{var}(\tilde{R})} - (A/2)V^{WS}_{mv} - (A/2)\frac{ER^2}{(A)^2 \,\text{var}(\tilde{R})} \geq$$
$$C^{SE}_{mv} + \frac{ER^2}{A \,\text{var}(\tilde{R})} - (A/2)V^{SE}_{mv} - (A/2)\frac{ER^2}{(A)^2 \,\text{var}(\tilde{R})} \tag{23}$$

All the terms involving ER cancel, meaning that equation (22) holds if and only if:

$$C_{mv}^{WS} - (A/2)V_{mv}^{WS} \geq C_{mv}^{SE} - (A/2)V_{mv}^{SE} \tag{24}$$

And, by equation (6), equation (24) holds if and only if

$$U\left(C_{mv}^{WS}, V_{mv}^{WS}\right) \geq U\left(C_{mv}^{SE}, V_{mv}^{SE}\right) \tag{25}$$

Thus, equation (21) holds if and only if equation (25) holds. In words, a consumer prefers the optimal point as a wage-earner to the optimal point being self-employed if and only if she prefers the minimum variance point in WS to the minimum variance point in SE. It is easy to see that if we replace "≥" with "=" in equations (21) to (25), we can also say that she is indifferent between wage-earning and self-employment if and only if she is indifferent between the minimum variance point in WS and the minimum variance point in SE.

The lower panel of Figure 2 illustrates the occupational choice decision graphically. As before, the characteristics of the wage and salary income stream upon which the graph is based are as listed in the first row or Table 5. The analogous characteristics for self-employment are in the second row, labeled $SE(1)$. For each occupation, the figure shows the indifference curve that is tangent to the feasible set, and the indifference curve that goes through the minimum variance point. Note that the vertical distance between the indifference curves associated with the best point for each occupation is the same as the vertical distance between the indifference curves that pass through the minimum variance points for each occupation. We can confirm this fact by appealing to equations (19) and (20). According to equation (19), the distance from V_{mv}^{WS} to V_{WS} is exactly the same as the distance from V_{mv}^{SE} to V_{SE}. (Both are equal to the negative of $ER^2/(A)^2 \operatorname{var}(\tilde{R})$.) Similarly, equation (20) tells us that the distance from C_{mv}^{WS} to C_{WS} and the distance from C_{mv}^{SE} to C_{SE} are exactly the same. (Both equal the negative of $ER^2/A \operatorname{var}(\tilde{R})$.) Therefore, if, as in the picture, the minimum variance point of WS lies on a higher indifference curve than the minimum variance point of SE, then the optimal point in WS must lie on a higher indifference curve as well.

3.1.6 Information Required to Measure C_{mv}^i and V_{mv}^i

We have now established that in this model, the minimum variance levels of lifetime income and the variance of income are critical determinants of occupational choice. How can we measure these two quantities? Clearly, we need to know the moments of asset returns ER and $\operatorname{var}(\tilde{R})$. And given the definitions of C_{mv}^i and V_{mv}^i(see equations (9) and (11) above, respectively),we require only two pieces of information about the income profile, $\text{var}(\tilde{y}_1^i)$ and β_y^i. These are precisely the magnitudes that we estimated in Section 2. In a multi-period model, the computation of the variance of income and its covariance with asset returns is slightly more complicated; we defer to Section 3.3 a detailed discussion of the relevant issues.

3.1.7 Risk aversion and occupational choice

With estimates of C_{mv}^i and V_{mv}^i for both self-employment (SE) and wage-earning (WS)in hand, we can investigate how risk aversion enters the choice between the occupations. From equation (24), our consumer is indifferent between them if and only if

$$A = A^* = 2 \times \frac{C_{mv}^{WS} - C_{mv}^{SE}}{V_{mv}^{WS} - V_{mv}^{SE}} \tag{26}$$

Intuitively, suppose that lifetime consumption in self-employment is very much larger than in wage earning (the numerator is a very negative number), while the variance of self-employment is just a bit larger than in wage-earning (the denominator is a slightly negative number). Then the right hand side of the equation is a large positive number, implying that the individual would be indifferent between wage-earning and self-employment only if he were extremely risk averse (that is, had a very large value of A). Note also that if the right hand side of equation (26) is negative, then equality can never hold, because A has to be positive. In this case, self-employment always dominates wage-earning because it has both higher lifetime consumption and lower lifetime variance.

The top panel of Figure 3 depicts two feasible sets, again using the parameter values from the first two rows of Table 5. To calculate A^*, draw a line between the minimum variance points of the two sets. This line is an indifference curve and we can use equation (7) to calculate the absolute risk aversion level associated with it. If a consumer's absolute risk aversion exceeds A^*, then the slope of her indifference curve exceeds that of an indifferent individual and she prefers WS, and vice versa if $A \leq A^*$. The bottom panel of Figure 3, which is based upon the row of Table 6 labeled $SE(2)$, portrays a situation in which only a risk-loving consumer would be indifferent between WS and SE – that is, $A^* < 0$.

3.1.8 Risk, reward and occupational choice

We can interpret our measures of lifetime consumption and its variance using the familiar language of risk and return. Assume without loss of generality that lifetime consumption and lifetime variance are both higher in self-employment. Then we can think of $\Delta C = C_{mv}^{SE} - C_{mv}^{WS}$ as the "reward" to self-employment and $\Delta V = V_{mv}^{SE} - V_{mv}^{WS}$

as the "risk." Substituting into equation (22), an individual chooses self-employment if the reward outweighs the (suitably scaled) risk, that is, if

$$\Delta C \geq (A/2)\Delta V \tag{27}$$

When relationship (27) holds with equality, it shows the increase in lifetime consumption required to compensate the individual for assuming a given incremental amount of lifetime risk.

Note that equation (26), which shows the critical value of risk aversion that would make the individual indifferent between the two occupations, and equation (27), which shows the critical value of the increment to consumption that would make her indifferent, are derived from the same underlying utility function. In effect, then, they are different ways of interpreting the same information. The perspective in equation (26) allows us to estimate the implicit degree of risk aversion that drives the choice between the two occupations. Hence, we can assess whether the degree of risk aversion implied by the model is "reasonable" given existing estimates in the literature. The perspective in equation (27) allows us to compare the *actual* difference between consumption in the two occupations with the minimum consumption required to induce individuals to assume the extra risk. Hence, it provides an explicit way of measuring the success of the model at explaining the data.

3.2 Different methods for calculating occupational risk

Suppose we observe the income profiles of one self-employed and one wage-and-salary worker, that both workers have the same level of lifetime income ($Y^{SE} = Y^{WS}$), but that income in self-employment is more volatile ($\text{var}(\tilde{y}_1^{SE}) > \text{var}(\tilde{y}_1^{WS})$). Suppose further that self-employment income has positive covariance with asset returns ($\beta_y^{SE} > 0$) and income as a wage-earner is uncorrelated with asset returns ($\beta_y^{WS} = 0$). Table 6 provides an example of such income profiles. The upper panel of Figure 4 plots the corresponding points in $C - V$ space.

Which occupation does the individual choose? Without the lifetime utility-maximizing framework, there are two natural methods:

1. *mean-variance method*: Compare the income means and variances associated with each point. In this case, both income streams allow for the same mean level of consumption but wage-earning gives lower variance of consumption. Thus, this method predicts that any risk-averse person prefers to be a wage-earner.

2. *CAPM method*: Find the plan with the higher market value, i.e., apply a higher discount rate to income whose covariance with risky assets is higher. Applying

the rule in this case, we note that both plans have the same present discounted value. But self-employment has higher covariance with the market portfolio, thus it is "riskier" and has a lower value. So any risk-averse individual prefers wage-earning.

What does our lifetime utility-maximizing method predict? Equations (19) and(20) allow us to use the income streams to calculate the minimum variance point in $C-V$ space and thus the feasible sets associated with wage-earning and self-employment. By equation (26), we can then calculate A^*, the absolute risk aversion level of a consumer indifferent between the two occupations. As indicated in Table 6, a consumer with absolute risk aversion (equal to/greater than/less than) 0.103 prefers (neither/SE/WS). (An absolute risk aversion coefficient of 0.103 corresponds to a local relative risk aversion coefficient of approximately 4.) This situation is depicted graphically in the lower panel of Figure 4, which plots the two feasible sets and the indifference curve of a consumer indifferent between the two occupations. In other words, the *more risk averse* the individual, the more likely she is to choose *self-employment* in spite of the fact that two standard measures of riskiness – the variance of the income stream and the covariance of the income stream with risky asset returns – suggest that self-employment is more risky.

Both the mean-variance and the CAPM approaches predict that no one would ever choose self-employment, so both are wrong. The mean-variance approach fails because it uses only the variance of income to measure risk and ignores the covariance of income with asset returns. The CAPM approach fails because it uses only the covariance of income with asset returns to measure risk and ignores the undiversifiable component of labor income. What makes the CAPM approach appropriate in standard finance applications is the assumption that for a given asset, an investor can diversify away all variation uncorrelated with the market – an assumption that is untenable in our setting.

3.3 Multi-period model

We now move to a multi-period framework. An investor chooses occupation i at time t. We assume that her decision is irreversible – once in occupation i, she must continue with occupation i until the end of time (date T). [10] As in the two-period model, preferences are defined over the stochastic consumption stream ($\{\tilde{c}_s\}_{s=t}^{T}$) and

[10] As we noted earlier, people can and do reverse their decisions. In terms of our theoretical model, *any* experiment with self-employment, no matter how short-lived, is irreversible in the sense that it permanently changes the characteristics of the future income stream.

represented by a time separable utility function with exponential period utility:

$$U\left(\{\tilde{c}_s\}_{s=t}^{T}\right) = \mathrm{E}_t\left[\sum_{s=t}^{T}(\delta_h)^{s-t}\left(\frac{-1}{A}\right)\exp\left(-Ac_s\right)\right]$$

Define the PDV operator, which takes the present discounted expected value of a random sequence,

$$\mathrm{PDV}_t\left(\{z_s\}_{s=t_1}^{t_2}\right) = \sum_{s=t_1}^{t_2}\frac{1}{R_0^{s-t_1}}\mathrm{E}\left(\tilde{z}_s\right).$$

Each period an investor must satisfy the period budget constraint:

$$\tilde{c}_t = \tilde{y}_t^i + \omega_{0,t-1}R_0 + \omega_{1,t-1}\tilde{R}_t - \omega_{0,t} - \omega_{1,t} \tag{28}$$

where $\omega_{0,t}$ and $\omega_{1,t}$ represent the holdings of the riskless and risky assets purchased at time t. As in the two-period model, the period budget constraint (equation (28)) implies a lifetime budget constraint:

$$C_t = \mathrm{PDV}_t(\{\tilde{c}_s\}_{s=t}^{T}) = \mathrm{PDV}_t(\{\tilde{y}_s^i + ER\omega_{1,s-1}\}_{s=t+1}^{T}) + W_t \tag{29}$$

where $W_t = \omega_{0,t-1}R_0 + \omega_{1,t-1}\tilde{R}_t$. [11]

To facilitate the use of this framework for analyzing the stylized facts from Section 2, we assume an income generating process of the same form that was estimated there. Using slightly different notation, this is:

$$\chi^i\left(L\right)\tilde{y}_t^i = \theta^i\left(L\right)\tilde{\eta}_t^i + \phi^i\left(L\right)\tilde{r}_t + d_t^i \tag{30}$$

where d_t^i is a deterministic constant, L is the lag operator, and χ^i, θ^i and ϕ^i are polynomial functions. We assume that the coefficient on the contemporaneous innovation in income is 1, that η_t^i and r_t are serially uncorrelated, and that they are uncorrelated with each other. A possible drawback to equation (30) is that the distribution of the level of income is highly skewed. However, specifying the equation in logs is unappealing in this context, because it is not uncommon for the self-employed to have negative incomes in a given year, and we do not want to discard these observations. [12] As a check on whether this skewness of income seriously biases our results, we present below calculations based on medians as well as means.

[11] In the two period model, we implicitly assume that initial asset holdings were zero, making $W_0 = 0$.

[12] To the extent that one excludes such observations, it lowers the variance and raises the mean of self-employment consumption. This would only reinforce our main finding from the next section–the increase in mean consumption that rewards increased variance in self-employment is much too large to be rationalized by conventional measures of risk aversion.

We show in the appendix (see Proposition 3) that under the maintained assumptions:

$$U\left(\{\tilde{c}_s\}_{s=t}^{T}\right) = U\left(C_t, V_t\right) = -\frac{1}{a_t A}\exp -a_t A\left(C_t - \frac{A}{2}V_t\right) \tag{31}$$

where $V_t = \text{PDV}_t(\{v_s/a_s\}_{s=t+1}^{T})$ $v_s = \text{var}_{s-1}(\tilde{c}_s)$ and $a_t = 1/\,\text{PDV}_t(\{1\}_{s=t}^{T})$ is an annuitization factor. In effect, we can write lifetime utility as a function of the present value of the expected lifetime consumption stream and the present value of its variance. In analogy with equation (6), then, equation (31) allows us to define indifference curves in $C_t - V_t$ space.

In analogy to the discussion leading up to equation (13) we can characterize the $C_t - V_t$ feasible set (see Proposition 2 in the appendix for details):

$$C_t = (ER/\sigma)\,\text{PDV}_t(\{(1/a_s)\}_{s=t+1}^{T})^{1/2}\left(V_t - V_{mv,t}^{i}\right)^{1/2} + W_t + C_{mv,t}^{i} \tag{32}$$

Here the point $(C_{mv,t}^{i}, V_{mv,t}^{i})$ represents the point in $C_t - V_t$ space with minimum V_t for an individual who chooses occupation i. Again, the two-period results carry through with appropriate re-definition of the key variables. The only difference is the presence of W_t on the right hand side; this reflects the fact that individuals can consume out of accumulated wealth.

The same reasoning as in the two-period case suggests that our consumer prefers the optimal feasible point in self-employment to the optimal feasible point as a wage-earner if and only if she prefers the minimum variance point in the set SE to the minimum variance point in the set WS. Hence, it is important to be able to calculate $(C_{mv,t}^{i}, V_{mv,t}^{i})$. To do so, suppose that our consumer chooses occupation i at time t. We show in the appendix (see equation (60) in Proposition 1) that with our assumptions:

$$\tilde{c}_s - \text{E}_{s-1}(\tilde{c}_s) = a_s\left(Y_t^{i} - \text{E}_{t-1}(Y_t^{i}) + \omega_{1,t-1}(\tilde{R}_s - \text{E}_{t-1}(\tilde{R}_t)\right) \tag{33}$$

where $Y_t^{i} = \text{PDV}_t(\{\tilde{y}_s^{i}\}_{s=t}^{T})$. In words, consumption responds to news about current and future labor income and to portfolio returns. Specifically, in a given year, the innovation to consumption is the annuity value of the sum of the innovations to the present discounted value of all future labor income $(Y_t^{i} - \text{E}_{s-1}(Y_t^{i}))$, and the innovation to the value of the portfolio $(\omega_{1,s-1}(\tilde{R}_s - \text{E}_{s-1}(\tilde{R}_s))$. Taking the variance of both sides of the equation,

$$v_s = a_s^2\,\text{var}_{s-1}(Y_s^{i} + \omega_{1,s-1}\tilde{R}_s)$$

Again in analogy to the two-period case, to minimize the variance of consumption, the individual sets

$$\omega_{1,s-1} = -B_{y,s}^{i}$$

where $B^i_{y,s} = \text{cov}_{s-1}(Y^i_s, \tilde{R}_s)/\,\text{var}_{s-1}(\tilde{R}_s)$. Let

$$v^i_{mv,s} = a^2_s\,\text{var}_{s-1}(\tilde{Y}^i_s - B^i_{y,s}\tilde{R}) \tag{34}$$

Thus:

$$V^i_{mv,t} = \text{PDV}_t(\{v^i_{mv,s}/a_s\}^T_{s=t+1}) \tag{35}$$
$$C^i_{mv,t} = \text{PDV}_t(\{\tilde{y}^i_s - B^i_{y,s}ER\}^T_{s=t+1}) \tag{36}$$

The lifetime minimum variance at a given point in time and the associated consumption level are just the suitably weighted present discounted values of the respective yearly minimum variance and associated consumption bundles.

3.3.1 Toward Implementing the Theory

To make this approach operational requires that we be able to measure $v^i_{mv,s}$ and $B^i_{y,s}$. Given that these parameters characterize consumption decisions, one might think that consumption data are required to estimate them. This would be a problem, because conventional panel data sets have inadequate measures of consumption or none at all. However, our model supplies a relationship between consumption and income, so that it is actually possible to estimate these parameters using the estimated income-generating functions from Section 2. Equation (30) has an associated moving average representation:

$$\tilde{y}^i_t = \psi^i\,(L)\,\tilde{\eta}^i_t + \beta^i_y\,(L)\,\tilde{r}_t + d^i_t \tag{37}$$

where $\psi^i_0 = 1$. [13] Equation (37) shows that both current income shocks and current asset returns shocks are useful for forecasting future income and thus the present discounted value of human capital. More explicitly, equation (37) implies that:

$$\tilde{y}^i_t - \text{E}_{t-1}(\tilde{y}^i_t) = \tilde{\eta}^i_t + \beta^i_{y,0}\tilde{r}_t, \tag{38}$$
$$\text{E}_t(\tilde{y}^i_{t+1}) - \text{E}_{t-1}(\tilde{y}^i_{t+1}) = \psi^i_1\tilde{\eta}^i_t + \beta^i_{y,1}\tilde{r}_t \tag{39}$$

and so on, where $\beta^i_{y,j}$ is the coefficient on the j^{th} lag of $\tilde{r}$. Intuitively, each period the individual uses new information contained in both earnings and asset returns to update her expectations about that period's income *and* income in all future periods.

Taking the present values of both sides of equations (38) and (39) and re-arranging, we obtain

$$Y^i_t - \text{E}_{t-1}(Y^i_t) = \Psi^i_t\eta_t + \text{PDV}_t(\{\beta^i_{y,j}\}^{T-t}_{j=0})\tilde{r}_t \tag{40}$$

[13] The coefficients on contemporaneous and lagged rates of return implied by the $\beta^i_y(L)$ function are the dynamic analogues of the beta of income with the risky asset, as defined in Section 3.1.3.

where $\Psi_t^i = \text{PDV}_t(\{\psi_j^i\}_0^{T-t})$. Taking the covariance of both sides of equation (40) with $\tilde{r}_t$, we obtain the covariance of Y_t^i and $\tilde{r}$ on the left hand side, and the variance of $\tilde{r}$ times $\text{PDV}_t(\{\beta_{y,j}^i\}_{j=0}^{T-t})$ on the right hand side. Taking advantage of the definition of $B_{y,s}^i$, this gives us

$$B_{y,s}^i = \text{PDV}_t(\{\beta_{y,j}^i\}_{j=0}^{T-t})$$

Next, substitute equation (40) into the definition of $v_{mv,s}^i$, equation (34). This yields

$$v_{mv,s}^i = \text{var}_{s-1}(a_s \Psi_s^i \tilde{\eta}^i) \tag{41}$$

So, all we need to characterize the feasible set is to estimate equation (30), because its parameters, in conjunction with the discount rate and the Sharpe ratio, determine the location and shape of the feasible set.

We now illustrate this computation with an example based on the specific functional form for the income generating equation that was estimated in Section 2–income follows a second order autoregressive process in first differences, and only a single lag of asset returns is useful for forecasting:

$$\tilde{y}_t - y_{t-1} = \tilde{\eta}_t + \alpha_0 + \chi_1 (y_{t-1} - y_{t-2}) + \chi_2 (y_{t-2} - y_{t-3}) + \phi_0 \tilde{r}_t + \phi_1 r_{t-1} \tag{42}$$

If we let:

$$\Gamma = \begin{bmatrix} 1+\chi_1 & \chi_2 - \chi_1 & -\chi_2 & 1 & \phi_0 & \phi_1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} Q_t = \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ \alpha_0 \\ r_t \\ r_{t-1} \\ \eta_t \end{bmatrix}$$

Then we can re-write equation (42) in matrix notation as:

$$Q_{t+1} = \Gamma Q_t \tag{43}$$

Equation (42) allows us to calculate the expectation of Q_s conditional on information at time t directly:

$$\text{E}_t(Q_s) = \Gamma^{s-t} Q_t \tag{44}$$

We can then calculate ψ_s the effect of a shock to income ($\tilde{\eta}_t$) on income at time s:

$$\psi_s = \Gamma^{s-t}(1, 7),$$

where the (1,7) denotes the element in the first row and seventh column of the matrix. This implies that

$$\Psi_s = \mathrm{PDV}_t(\{\Gamma^{s-t}(1,7)\}_{s=t}^{T}) \tag{45}$$

Similarly, $\beta_{y,s}$, the effect of a shock to returns ($\tilde{r}_t$) today on income at time s, is

$$\beta_{y,s} = \Gamma^{s-t}(1,5) \tag{46}$$

and

$$B_{y,s} = \mathrm{PDV}_t(\{\Gamma^{s-t}(1,5)\}_{s=t}^{T}) \tag{47}$$

Now we add back the i superscripts. If we assume that $\tilde{\eta}_t^i$ is homoskedastic, then equations (35) and (41) imply that:

$$V_{mv,t}^{i} = \mathrm{var}(\tilde{\eta}^i)\,\mathrm{PDV}_t(\{(a_s\Psi_s^i)^2\}_{s=t+1}^{T}) \tag{48}$$

Hence, all the information we require comes from estimated parameters of equation (42) – $\mathrm{var}(\tilde{\eta}^i)$ falls out directly from estimating the equation (it is the residual variance), and Ψ_s^i is the function of the parameters described in equation (45). Similarly (36) implies that:

$$C_{mv,t}^{i} = Y_t^i - ER \times \mathrm{PDV}_t(\{B_{y,s}^i\}_{s=t+1}^{T}) \tag{49}$$

Again, given an initial value of income, we can use equation (42) to project income forward and hence estimate estimate Y_s^i. Further, $B_{y,s}^i$ depends on the parameters of equation (42) via equation (47). Hence, the income generating process provides all the information we need to characterize the minimum variance consumption point, as asserted at the beginning of this section.

4 Implications for the Choice Between Wage-Earning and Self-Employment

Given the theory from the previous section, we are now in a position to interpret the empirical findings from Section 2 which were summarized in Table 2. The calculations that our critical according to the theory are reported in Table 7 for the sample as a whole and by educational class. The column headed Y_t^i is the present discounted value of lifetime income. This calculation is done for a “typical” thirty year old. Specifically, we take the mean income for a person between 28 and 32, use the polynomial for age in the income equation to calculate the associated expected income profile (assuming that she retires at age 65 and dies at age 75), and discount this stream at a rate of return of 2.5 percent.

The next column shows the present discounted value of $B^i_{y,t}$, which is calculated using the β's in Table 2, substituting them into equation (47), and then taking the present value of the resulting series. The third column shows the minimum variance consumption level, computed by substituting into equation (49). This simply requires taking the figure for Y^i_t from the first column, and subtracting from it .05 (our assumed risk premium) times the present discounted value of the $B^i_{y,t}$ from the second column. Ψ^i_t is found by substituting into equation (45), evaluated at age 30, and $V^i_{mv,t}$ by substituting into equation (48).

The figures in Table 7 indicate that, for both educational classes, the minimum variance level of lifetime consumption is higher in self-employment than in wage-earning, but so is the variance of lifetime consumption. A critical question is whether, given what is know about risk aversion, the extra consumption available in self-employment is about the right magnitude to compensate for the additional variance. Put another way, what level of risk aversion would be required for an individual to be just indifferent between the consumption-variance bundles available to her? This is precisely the question answered by equation (26), which shows how to compute the critical value of absolute risk aversion. We substitute the figures from Table 7 into equation (26), and compute the (locally) equivalent levels of constant relative risk aversion by multiplying by average income.[14]

The results are reported in the column headed RRA. The figures of 128.9 and 54.6 for the lower and higher educational classes, respectively, are many times conventional estimates of relative risk aversion, which are generally in the range of 1 to 3. At face value, these figures imply that only a pathologically risk- averse person would turn down the extra rewards available to the self-employed. What are we to make of this? Our conclusion is that it is very hard to reconcile cross-sectional differences in the choice between self-employment and wage-earning with conventional estimates of risk aversion. This is contrary to the dominant strain in the literature, which views a willingness to accept risk as the key determinant of the decision to become self-employed.[15]

As noted above, another way to interpret our results is in terms of the implied increment to consumption just needed to compensate the individual for the additional risk associated with self-employment. (See the discussion surrounding equation (27).)

[14] The average income figure is just the average of the incomes associated with the two occupational modes from Table 3.

[15] Interestingly, on the basis of U.S. survey evidence, Barsky, Juster, Kimball and Shapiro (1997) show that the self-employed are not significantly more risk tolerant than wage earners. However, Cramer *et al.'s* (2001) analysis of Dutch survey data finds a negative relationship between risk aversion and the probability of ever having been self-employed.

For purposes of this computation we assume a value for RRA of 3. The results, reported in the column of Table 7 labeled ΔC, are expressed as a fraction of the *actual* consumption differential. Thus, for example, the figure of 1.17 in the "no more than high school" row means that, for these individuals, the model "explains" only 1.17 percent of the difference between the consumption levels of wage-earners and the self-employed. Presumably, factors other than those considered by the model must be at work here; we return to this notion later.

We computed the implied values of RRA and ΔC for all specifications of the model. Some of the results are reported in Table 8 (We report only the results, not the intermediate calculations, for the sake of brevity.) The first bank of numbers basically reproduces for reference the results from the canonical model (Table 7). In addition, we show calculations of RRA and ΔC using median rather than mean consumption. The calculations based on means and medians are qualitatively very similar; hence, the skewness in the distribution of consumption does not appear to be seriously affecting our results. The next two banks of figures use the results from the pseudo-IV and "irrevocable" estimates of the income generating equations, respectively. Again, the basic findings hold: the implied degrees of relative risk aversion are implausibly high, and the model explains only a very small proportion of the actual differential between consumption levels.

The bottom half of the table is motivated by the fact that, for some education groups, the estimates of β_y^i, the covariance between income and asset returns, are imprecisely estimated. It seems worthwhile to investigate whether our results are being driven by point estimates that, from a statistical point of view, could just as well be zero. Therefore, we re-estimated RRA and ΔC simply setting all the β_y^i equal to zero. The figures suggest that our results are not highly sensitive to the estimates of the covariance between income and asset returns. Thus, our findings with respect to the limitations of an approach that puts risk aversion at the center of the occupational choice decision are not an artifact of the particular "betas" that we estimate.

5 Conclusion

Both economists and laypeople place attitudes toward risk at the center of the self-employment choice. To assess the validity of this view, we have constructed a model of the choice between wage-earning and self-employment and calibrated it using panel data. The model is based on a modern framework for analyzing lifetime utility maximization under uncertainty, and takes seriously the notion that risk aversion is the driving factor in the decision to become self-employed. Thus, for example, we do not take into account possible unobservable differences among individuals in their opportunity sets. These might include differences in the talents required for success in self-employment (perhaps arising from differences in family background), or differences in wealth (which can be important in the presence of capital market imperfections and substantial start-up expenses).[16] In effect, our goal is to see how well an approach that ignores such considerations and focuses on risk aversion does at explaining the data.

The answer is that such a tack does not do very well at all. We find that in order to rationalize the choices we observe in the data, people would have to be pathologically risk averse. Hence, risk aversion alone does not seem to be the predominant consideration driving the decision. Indeed, we find evidence that it is not even an important consideration–the amount of additional lifetime consumption that individuals would require to accept the incremental risk associated with self-employment is only a small fraction of the actual consumption differential.

The failure of the risk-aversion explanation puts the spotlight on the unobservable variables just discussed, and these surely deserve consideration in future research. However, some more fundamental problems with the model could also be involved. One such problem is that people might simply fail to evaluate correctly the risks associated with self-employment. The literature includes a number of famous examples of behavior toward risk that have puzzled economists, including the absence of any risky assets in many households' portfolios (see, for example, Poterba and Samwick (1995)), and a premium on the return to owning equity that is much higher than can be rationalized by conventional measures of risk aversion (Mehra and Prescott (1985)). The self-employment choice may turn out to be another one of these puzzles.

[16] See Hout and Rosen (2000) on the impact of family background in the self-employment decision, and Evans and Jovanovic (1989) and Gentry and Hubbard(2000) on the role of liquidity constraints.

References

[1] Barsky, Robert, F. Thomas Juster, Miles Kimball and Matthew D. Shapiro. 1997. Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *Quarterly Journal of Economics* 112(2):537-79.

[2] Blanchflower, D.G. and A.J. Oswald. 1998. What makes an entrepreneur? *Journal of Labor Economics* 16 (1):26-60.

[3] Borjas, George and Stephen Bronars. 1989. Consumer discrimination and self-employment. *Journal of Political Economy* 97:581-605.

[4] Brockhaus, Sr., Robert H. 1980. Risk taking propensity of entrepreneurs. *Academy of Management Journal* 23(3):509-20.

[5] Cramer, J.S., J. Hartog, N. Jonker, and C.M. Van Praag. 2001. Low risk aversion encourages the choice for entrepreneurship: An empirical test of a truism. University of Amsterdam. Working paper.

[6] Davis, Steven and Paul Willen. 2000a. Using financial assets to hedge labor income risks: Estimating the benefits. University of Chicago. Working paper.

[7] —. 2000b. Occupation-level income shocks and asset returns: Their covariance and implications for portfolio choice. NBER Working Paper No. 7905.

[8] Deaton, Angus. 1991. Saving and liquidity constraints. *Econometrica* 59:1221-48.

[9] Evans, David S. and Boyan Jovanovic. 1989. An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy* 97:808-27.

[10] Fairlie, Robert and Bruce Meyer. 1996. Ethnic and racial self-employment differences and possible explanations. *Journal of Human Resources* 31(4):757-93.

[11] Gentry, William M. and R. Glenn Hubbard. 2000. Entrepreneurship and household saving. NBER working paper 7894.

[12] Hamilton, Barton, H. 2000. Does entrepreneurship pay? An empirical analysis of the returns to self-employment. *Journal of Political Economy* 108:604-31.

[13] Heaton, John and Deborah Lucas. 2000. Portfolio choice and asset prices: The importance of entrepreneurial risk. *Journal of Finance* 55:1163-98.

[14] Hout, Michael and Harvey S. Rosen. 2000. Self-employment, family background, and race. *Journal of Human Resources* 35:670-92.

[15] Hurst, Erik and Annamaria Lusardi. 2001. Liquidity constraints, wealth accumulation and entrepreneurship. University of Chicago. Working paper.

[16] Kanbur, S.M. 1979. Of risk taking and the personal distribution of income. *Journal of Political Economy* 87(4):769-797.

[17] Kihlstrom, Richard E. and Jean-Jacques Laffont. 1979. *Journal of Political Economy* 87(4):719-748.

[18] Knight, Frank H. 1971. */Risk, Uncertainty and Profit* (ed. G.J. Stigler), Chicago: University of Chicago Press. (first edition 1921).

[19] MaCurdy, Thomas. 1982. The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics* 18:83-114.

[20] Mehra, Rajnish and Edward C. Prescott. 1985. The equity premium: A puzzle. *Journal of Monetary Economics* 15:145-61.

[21] Moskowitz, Tobias and Annette Vissing-Jorgensen. 2001. The private equity premium puzzle. University of Chicago. Working paper.

[22] Newman, Andrew F. 1995. Risk-bearing and ‘Knightian’ entrepreneurship. Columbia University. Working paper.

[23] Poterba, James and Andrew Samwick. 1995. Stock ownership patterns, stock market fluctuations and consumption. *Brookings Papers on Economic Activity* 2:295-372.

[24] Quadrini, Vincenzo. 1999. The importance of entrepreneurship for wealth concentration and mobility. *Review of Income and Wealth* 45(1):1-19.

Appendix

For all the following proofs, we consider a consumer who chooses occupation i. For simplicity, we suppress the i superscripts unless necessary.

Condition 1 *Agents have exponential utility,*

$$U(\mathbf{C}) = \mathrm{E}_t\left[\sum_{t=t}^{T} (\delta_h)^t \left(\frac{-1}{A}\right) \exp(-Ac_t)\right]$$

where A is the coefficient of absolute risk aversion.

Condition 2 *The distribution of risky asset returns is non-stochastic. I.e.* $\mathrm{E}_s(\tilde{r}_t) = R_0 + ER \ \ \forall s < t$ *and* $\mathrm{var}_s(\tilde{R}_t) = \mathrm{var}(\tilde{R}) \ \ \forall s < t$.

Condition 3 *Income follows an ARIMA process:*

$$\chi(L)\, y_t = \theta(L)\, \eta_t + \phi(L)\, r_t + d_t \tag{50}$$

where d_t is a deterministic constant. Assume that η_t and r_t are jointly normally distributed and iid over time. Equation (50) has associated $\mathrm{MA}(\infty)$ *representation:*

$$y_t = \psi(L)\, \eta_t + \beta_y(L)\, r_t + d_t$$

Lemma 1 *Let $Y_t = \mathrm{PDV}_t(\{\tilde{y}_s\}_{s=t}^T)$. Under Condition 3*

$$Y_t - \mathrm{E}_t(Y_t) = \mathrm{PDV}_t(\{\psi_j\}_0^{T-t})\eta_t + \mathrm{PDV}_t(\{\beta_{y,j}\}_{j=0}^{T-t})r_t$$

Let $\Psi_t = \mathrm{PDV}_t(\{\psi_j\}_0^{T-t})$ and $B_{y,t} = \mathrm{PDV}_t(\{\beta_{y,j}\}_{j=0}^{T-t})$ Then

$$Y_t - \mathrm{E}_t(Y_t) = \Psi_t \eta_t + B_{y,t} r_t$$

Lemma 2 *Assume that $\tilde{c}_{t+1}$ and $\tilde{R}$ are jointly normally distributed and that Conditions 1 and 2 hold. Then:*

$$ER = A\,\mathrm{cov}(\tilde{c}_{t+1}, \tilde{R}) \tag{51}$$

$$\mathrm{E}_t(c_{t+1}) - c_t = \frac{A}{2}\,\mathrm{var}_t(c_{t+1}) - \frac{1}{A}\ln R_0\delta \tag{52}$$

Proof: The intertemporal Euler equations are:

$$u_c(c_t) = \delta R_0\, \mathrm{E}_t(u_c(c_{t+1})) \tag{53}$$

$$u_c(c_t) = \delta\, \mathrm{E}_t(u_c(c_{t+1})\tilde{R}) \tag{54}$$

We use the definition of covariance and Stein's Lemma to transform equation (54) into:

$$u_c(c_t) = \delta\, \mathrm{E}_t(u_c(c_{t+1}))\, \mathrm{E}_t(\tilde{R}) + \delta\, \mathrm{E}_t(u_{cc}(c_{t+1}))\, \mathrm{cov}_t(\tilde{c}_{t+1}, \tilde{R}) \tag{55}$$

To get equation (51), divide equation (55) by equation (53) and re-arrange. Since we assume that $\tilde{c}_{t+1}$ is normally distibuted, we can re-write equation (53):

$$\exp -Ac_t = R_0 \delta \exp \left[-A \, \mathrm{E}_t(c_{t+1}) + (A/2) \, \mathrm{var}_t(c_{t+1}) \right] \tag{56}$$

To get equation (52), take logs of both sides of equation (56) and divide through by $-A$ □

Proposition 1 (Individual optimization) *Assume the consumer chooses occupation i. Given conditions 1, 2 and 3:*

1. *Risky asset holdings in the optimal portfolio are given by*

$$\omega_t = \frac{1}{A_{t+1}} \frac{ER}{\mathrm{var}(\tilde{R})} - B_{y,t+1} \tag{57}$$

where $A_t = a_t A$.

2. *Consumption is:*

$$c_t = a_t \left(C_t - \frac{A}{2} V_t \right) \tag{58}$$

where

$$\begin{aligned} C_t &= \mathrm{PDV}_t(\{\tilde{c}_s\}_{s=t}^{T}) \\ &= Y_t + \omega_{0,t-1} R_0 + \tilde{r}_t \omega_{t-1} \, \mathrm{PDV}_t(\{ER\omega_s\}_{s=t}^{T-1}) \\ V_t &= \mathrm{PDV}_t(\{\mathrm{var}_{s-1}(\tilde{c}_s)/a_s\}_{s=t+1}^{T}) \\ &= \mathrm{PDV}_t(\{a_s \, \mathrm{var}_{s-1}(\tilde{\Psi}_s \tilde{\eta}_s) + a_s \, \mathrm{var}_{s-1}(\tilde{r}_s \, (\omega_{s-1} + B_{y,s}))\}_{s=t+1}^{T}) \end{aligned} \tag{59}$$

3. *Consumption innovations are:*

$$\tilde{c}_t - \mathrm{E}_{t-1}(\tilde{c}_t) = a_t \Psi_t \tilde{\eta}_t + a_t \tilde{r}_t \, (\omega_{t-1} + B_{y,t}) \tag{60}$$

Proof: At time $T-1$,

$$\tilde{c}_T = \tilde{y}_T + \omega_{0,T-1} R_0 + \omega_{T-1} \tilde{r}_T \tag{61}$$

Condition 3 and equation (61) imply that consumption and asset returns are jointly normal. Thus Lemma 2 applies and:

$$\mathrm{E}_{T-1}(\tilde{c}_T) = c_{T-1} + \frac{A}{2} \, \mathrm{var}_{T-1}(\tilde{c}_T) \tag{62}$$

By definition:

$$C_{T-1} = c_{T-1} + \frac{1}{R_0} \, \mathrm{E}_{T-1}(\tilde{c}_T) \tag{63}$$

Combining equations (62) and (63) yields:

$$C_{T-1} = c_{T-1} + \frac{1}{R_0}\left(c_{T-1} + \frac{A}{2}\,\text{var}_{T-1}(\tilde{c}_T)\right) \tag{64}$$

By the intertemporal budget constraint:

$$C_{T-1} = \text{PDV}_{T-1}(\{\tilde{c}_s\}_{s=T-1}^{T}) = Y_{T-1} + \omega_{0,T-2}R_0 + \tilde{r}_{T-1}\omega_{T-2} + \text{PDV}_{T-1}(\{ER\omega_s\}_{s=T-1}^{T-1}) \tag{65}$$

Combining equations (64) and (65) and rearranging gives

$$c_{T-1} = a_{T-1}\left[Y_{T-1} + \omega_{0,T-2}R_0 + \tilde{r}_{T-1}\omega_{T-2}\right] + a_{T-1}\left[\text{PDV}_{T-1}(\{ER\omega_s\}_{s=T-1}^{T-1}) - \frac{1}{R_0}\left(\frac{A}{2}\,\text{var}_{T-1}(\tilde{c}_T)\right)\right] \tag{66}$$

To calculate portfolio holdings at $T-1$, we use the time T period budget constraint (equation (61)) to replace consumption in equation (27):

$$\begin{aligned}\text{cov}_{T-1}(\tilde{c}_T, \tilde{r}_T) &= \text{cov}_{T-1}(\tilde{y}_T + \omega_{T-1}\tilde{r}_T, \tilde{r}_T)\\ &= \text{cov}_{T-1}(\tilde{\eta}_T + (\beta_{y,T} + \omega_{T-1})\,\tilde{r}_T, \tilde{r}_T)\\ &= \text{var}(\tilde{R})\,(\beta_{y,T} + \omega_{T-1})\end{aligned} \tag{67}$$

To get portfolio holdings combine equation (67) and equation (27) from Lemma 2,

$$\omega_{T-1} = \frac{ER}{A\,\text{var}(\tilde{R})} - \beta_{y,T} \tag{68}$$

In addition, we can solve for the variance of consumption using equation (68),

$$\text{var}_{T-1}(\tilde{c}_T) = \text{var}_{T-1}(\tilde{\eta}) + \frac{1}{(A)^2}\frac{ER^2}{\text{var}(\tilde{r})} \tag{69}$$

Now we consider the $T-2$ decision problem. Consider equation (66). By equations (68) and (69) respectively, $\text{PDV}_{T-1}(\{ER\omega_{s-1}\}_{s=T}^{T})$ and $\text{var}_{T-1}(\tilde{c}_T)$ are in the time $T-2$ information set. Thus:

$$\tilde{c}_{T-1} - \text{E}_{T-2}(\tilde{c}_{T-1}) = a_{T-1}\left(\tilde{Y}_{T-1} - \text{E}_{T-2}(\tilde{Y}_{T-1}) + \tilde{r}_{T-1}\omega_{T-2}\right) \tag{70}$$

Equation (70) and Lemma 1 imply that:

$$\tilde{c}_{T-1} - \text{E}_{T-2}(\tilde{c}_{T-1}) = a_{T-1}\Psi_{T-1}\tilde{\eta}_{T-1} + a_{T-1}\tilde{r}_{T-1}\,(\omega_{T-2} + B_{y,T-1}) \tag{71}$$

Condition 3 and equation (71) imply that consumption and asset returns are jointly normal and we can again apply Lemma 2. The derivation c_{T-2} and ω_{T-2} is identical to the one we just used to derive c_{T-1} and ω_{T-1}. To get risky asset holdings, use equation (51) from Lemma 2:

$$\omega_{T-2} = \frac{ER}{a_{T-1} A \operatorname{var}(\tilde{r})} - B_{y,T-1} \tag{72}$$

Equation (72) and equation (71) imply that:

$$\operatorname{var}_{T-2}(\tilde{c}_{T-1}) = a_{T-1}^2 \Psi_{T-1}^2 \operatorname{var}(\tilde{\eta}) + \frac{1}{(A)^2} \frac{ER^2}{\operatorname{var}(\tilde{R})} \tag{73}$$

To get consumption at time $T-2$, use equation (52) from Lemma 2 to calculate the value of lifetime consumption:

$$C_{T-2} = \left(1 + \frac{1}{R_0} + \frac{1}{R_0^2}\right) c_{T-2} + \mathrm{PDV}_{T-2}(\{A \operatorname{var}_{s-1}(\tilde{c}_s)/2a_s\}_{s=T-1}^{T}) \tag{74}$$

The intertemporal budget constraint (equation (29)) implies that:

$$C_{T-2} = Y_{T-2} + \omega_{0,T-3} R_0 + \tilde{r}_{T-1} \omega_{T-3} + \mathrm{PDV}_{T-2}(\{ER\omega_s\}_{s=T-2}^{T-1}) \tag{75}$$

Combining equations (71), (73) and (75) and rearranging yields:

$$\begin{aligned} \frac{c_{T-2}}{a_{T-2}} &= Y_{T-2} + \omega_{0,T-3} R_0 + \tilde{r}_{T-1} \omega_{T-3} + \mathrm{PDV}_{T-2}(\{ER\omega_s\}_{s=T-2}^{T-1}) \\ &\quad - \mathrm{PDV}_{T-2}(\{A_s \operatorname{var}_{s-1}(\Psi_s \tilde{\eta}) + \frac{ER^2}{A_s \operatorname{var}(\tilde{R})}\}_{s=T-1}^{T}) \end{aligned} \tag{76}$$

To solve for consumption for arbitrary time t, we continue working backwards and get

$$\begin{aligned} \frac{c_t}{a_t} &= Y_t + \omega_{0,t-1} R_0 + \tilde{r}_t \omega_{t-1} + \mathrm{PDV}_t(\{ER\omega_s\}_{s=t}^{T-1}) \\ &\quad - \mathrm{PDV}_t(\{A_s \operatorname{var}_{s-1}(\Psi_s \tilde{\eta}) + \frac{ER^2}{A_s \operatorname{var}(\tilde{R})}\}_{s=t+1}^{T}) \end{aligned}$$

$$\omega_{t-1} = \frac{\frac{ER}{\operatorname{var}(\tilde{r})}}{A_t} - B_{y,t}$$

$$\tilde{c}_t - \mathrm{E}_{t-1}(\tilde{c}_t) = a_t \Psi_t \tilde{\eta}_t + a_t \tilde{r}_t (\omega_{t-1} + B_{y,t})$$

□

Proposition 2 *Under Conditions 1, 2 and 3,*

$$C_t = (ER/\sigma)\,\mathrm{PDV}_t(\{(1/a_s)\}_{s=t+1}^T)^{1/2}\left(V_t - V_{mv,t}^i\right)^{1/2} + W_t + C_{mv,t}^i \tag{77}$$

Proof: First, substitute optimal portfolio choice (equation (57) into the intertemporal budget constraint (equation (29)),

$$C_t = \mathrm{PDV}_t(\{\tilde{y}_s - ERB_{y,s-1}\}_{s=t+1}^T) + \mathrm{PDV}_t(\{ER^2/A_t\,\mathrm{var}(\tilde{R})\}_{s=t+1}^T) + W_t \tag{78}$$

We can then characterize C_t as a function of $C_{mv,t}$, risk aversion and the distribution of returns. The definition of $C_{mv,t}$ (equation (35)), the definition of A_t and equation (78) yields:

$$C_t = C_{mv,t} + ER^2/A^2\,\mathrm{var}(\tilde{R})\,\mathrm{PDV}_t(\{1/a_s\}_{s=t+1}^T) + W_t \tag{79}$$

To characterize V_t, we can use the definition of V_t (see the text after equation (31)), equation (60), Lemma 1 and optimal portfolio choice (equation (57)) to get:

$$V_t = \mathrm{PDV}_t(\{v_{mv,s}\}_{s=t+1}^T) + \frac{ER^2}{A\,\mathrm{var}(\tilde{R})}\,\mathrm{PDV}_t(\{1/a_s\}_{s=t+1}^T) \tag{80}$$

$V_{mv,t} = \mathrm{PDV}_t(\{v_{mv,s}\}_{s=t+1}^T)$, by definition. We now combine equation (79) and equation (80) to get equation (77) □

Proposition 3

$$U\left(\{\tilde{c}_s\}_{s=t}^T\right) = -\frac{1}{a_t A}\exp -a_t A\left(C_t - \frac{A}{2}V_t\right) \tag{81}$$

Proof: Equation (52) in Lemma 2 shows that: $\exp(-Ac_t) = \delta R_{0,t+1}\,\mathrm{E}\left(\exp\left(-A\tilde{c}_{t+1}\right)\right)$ which implies that:

$$-\frac{(\delta)^\tau}{A}\,\mathrm{E}\left(\exp\left(-A\tilde{c}_\tau\right)\right) = -\frac{1}{A}\frac{1}{R_0^\tau}\exp\left(-Ac_t\right) \tag{82}$$

Equation (82) implies that

$$U\left(\{\tilde{c}_s\}_{s=t}^T\right) = -\frac{1}{A}\,\mathrm{PDV}_t\left(\{\mathbf{1}\}_{s=t}^T\right)\exp -Ac_t \tag{83}$$

To get equation (81), substitute equation (58) into equation (83) □

Table 1: Incidence of self-employment. This table shows the distribution of years of self-employment among different groups in the sample. The first figure in each cell is the number of people in that cell; the second number is the percentage of people in the associated column.

	At most High School	Some college or more	Men	Women
Total	10052	4981	10422	4611
As % of total	100.0	100.0	100.0	100.0
Never	8203	3662	7750	4115
As % of total	81.6	73.5	74.4	89.2
At least a year	1849	1319	2672	496
As % of total	18.4	26.5	25.6	10.8
At least two years	1066	860	1715	211
As % of total	10.6	17.3	16.5	4.6
At least five years	447	470	866	51
As % of total	4.4	9.4	8.3	1.1
At least ten years	206	245	440	11
As % of total	2.0	4.9	4.2	0.2
Exactly a year	783	459	957	285
As % of total	7.8	9.2	9.2	6.2

Table 2: Estimates of the household income generating model. All equations include individual fixed effects. A person is considered self-employed if he/she reports "self" as employer for five years or more. The p-value below R^2 is for an F-test of the joint significance of all the regressors. F_{NYSE} is an F-statistic for the restriction that the coefficients on current and lagged returns are both zero. Means, medians, and standard deviations are measured in thousands of 1996 dollars.

Education	SE status	Mean (median)	std(η) (std(η)/E($\tilde{y}$))	ψ_1 (se)	ψ_2 (se)	β_{NYSE} (se)	$\beta_{NYSE_{lag}}$ (se)	R^2_{NYSE} (p-val)	R^2 (p-val)	Nobs
all	all	43.3	20.6	-0.40	-0.17	-0.8	-0.5	1.95	0.13	90140
		(37.8)	(0.48)	(0.00)	(0.00)	(0.5)	(0.5)	(0.14)	(0.00)	
	Wage and	29.9	10.4	-0.33	-0.20	-0.7	-0.7	5.11	0.11	52553
No more than	Salary	(27.5)	(0.35)	(0.00)	(0.00)	(0.3)	(0.3)	(0.01)	(0.00)	
high school	Self	47.0	25.3	-0.47	-0.20	-4.7	2.0	3.90	0.18	6713
	Employed	(40.4)	(0.54)	(0.01)	(0.01)	(1.9)	(1.9)	(0.02)	(0.00)	
	Wage and	54.2	15.8	-0.38	-0.22	1.6	0.3	2.29	0.13	24060
Some college	Salary	(50.7)	(0.29)	(0.01)	(0.01)	(0.8)	(0.7)	(0.10)	(0.00)	
or more	Self	80.3	60.0	-0.40	-0.14	-5.3	-3.2	0.61	0.13	6121
	Employed	(62.3)	(0.75)	(0.01)	(0.01)	(5.2)	(5.1)	(0.54)	(0.00)	
	Wage and	39.5	12.4	-0.36	-0.21	-0.1	-0.5	1.68	0.12	76613
All educ.	Salary	(35.8)	(0.31)	(0.00)	(0.00)	(0.3)	(0.3)	(0.19)	(0.00)	
groups	Self	65.2	45.3	-0.41	-0.15	-4.9	-0.2	1.75	0.14	12834
	Employed	(50.0)	(0.70)	(0.01)	(0.01)	(2.6)	(2.6)	(0.17)	(0.00)	

Table 3: Income and its variability: Regression-corrected summary measures

		No more than high school		Some college or more		All educ groups	
		WS	SE	WS	SE	WS	SE
Household income							
5 year rule	Mean(median)	29.9 (27.5)	47.0 (40.4)	54.2 (50.7)	80.3 (62.3)	39.5 (35.8)	65.2 (50.0)
	$std_\eta(std_\eta\,/\,E\,\tilde{y})$	10.4 (0.35)	25.3 (0.54)	15.8 (0.29)	60.0 (0.75)	12.4 (0.31)	45.3 (0.70)
10 year rule	Mean(median)	30.5 (28.1)	52.0 (43.1)	56.2 (51.4)	81.5 (63.2)	40.9 (36.4)	69.0 (52.9)
	$std_\eta(std_\eta\,/\,E\,\tilde{y})$	11.0 (0.36)	29.1 (0.56)	18.0 (0.32)	72.2 (0.89)	13.7 (0.33)	54.6 (0.79)
Pseudo-IV	Mean(median)	34.3 (30.2)	50.2 (44.1)	68.5 (61.8)	83.7 (71.0)	44.3 (38.0)	68.3 (56.8)
	$std_\eta(std_\eta\,/\,E\,\tilde{y})$	12.5 (0.37)	23.8 (0.47)	34.3 (0.50)	41.1 (0.49)	21.3 (0.48)	34.2 (0.50)
Irrevocable	Mean(median)	29.2 (26.6)	38.0 (33.9)	54.7 (51.3)	65.9 (54.4)	38.9 (35.2)	51.7 (42.6)
	$std_\eta(std_\eta\,/\,E\,\tilde{y})$	9.9 (0.34)	18.5 (0.49)	15.0 (0.27)	44.6 (0.68)	11.7 (0.30)	31.8 (0.62)
Head's earnings							
5 year rule	Mean(median)	21.9 (20.5)	28.0 (23.3)	40.6 (37.9)	53.7 (38.3)	29.3 (26.9)	42.0 (30.6)
	$std_\eta(std_\eta\,/\,E\,\tilde{y})$	7.2 (0.33)	15.5 (0.56)	11.1 (0.27)	36.7 (0.68)	8.6 (0.29)	27.9 (0.66)
10 year rule	Mean(median)	22.1 (20.5)	29.9 (24.8)	41.6 (38.1)	54.2 (36.9)	30.0 (27.1)	43.9 (31.0)
	$std_\eta(std_\eta\,/\,E\,\tilde{y})$	7.6 (0.35)	17.0 (0.57)	13.0 (0.31)	42.2 (0.78)	9.7 (0.32)	32.1 (0.73)
Pseudo-IV	Mean(median)	22.0 (19.5)	27.9 (24.7)	48.4 (43.4)	51.1 (38.5)	29.7 (25.3)	40.5 (31.2)
	$std_\eta(std_\eta\,/\,E\,\tilde{y})$	8.2 (0.37)	13.8 (0.49)	19.7 (0.41)	30.6 (0.60)	12.7 (0.43)	24.3 (0.60)
Irrevocable	Mean(median)	21.8 (20.4)	24.6 (21.9)	41.1 (38.3)	46.0 (36.7)	29.1 (27.2)	35.1 (28.1)
	$std_\eta(std_\eta\,/\,E\,\tilde{y})$	6.6 (0.30)	12.1 (0.49)	10.0 (0.24)	27.9 (0.61)	7.8 (0.27)	20.2 (0.58)

Table 4: Estimates of the household income generating model using "pseudo-instrumental variables." All equations include individual fixed effects. A person is considered self-employed if he/she reports ever having been self-employed before the age of 35; the equations are estimated using only individuals who are 35 years of age and older. The p-value below R^2 is for an F-test of the joint significance of all the regressors. F_{NYSE} is an F-statistic for the restriction that the coefficients on current and lagged returns are both zero. Means, medians, and standard deviations are measured in thousands of 1996 dollars.

Education	SE status	Mean (median)	std(η) (std(η)/E($\tilde{y}$))	ψ_1 (se)	ψ_2 (se)	β_{NYSE} (se)	$\beta_{NYSE_{lag}}$ (se)	R^2_{NYSE} (p-val)	R^2 (p-val)	Nobs
all	all	46.6	23.0	-0.40	-0.16	-1.2	-0.7	2.29	0.13	64290
		(39.4)	(0.49)	(0.00)	(0.00)	(0.6)	(0.6)	(0.10)	(0.00)	
	Wage and	34.3	12.5	-0.36	-0.19	-0.9	-0.5	3.50	0.12	40606
No more than	Salary	(30.2)	(0.37)	(0.00)	(0.01)	(0.4)	(0.4)	(0.03)	(0.00)	
high school	Self	50.2	23.8	-0.49	-0.20	-7.7	0.6	2.99	0.20	3068
	Employed	(44.1)	(0.47)	(0.02)	(0.02)	(3.5)	(3.3)	(0.05)	(0.00)	
	Wage and	68.5	34.3	-0.41	-0.15	-1.7	-0.1	0.42	0.13	16566
Some college	Salary	(61.8)	(0.50)	(0.01)	(0.01)	(1.9)	(1.8)	(0.66)	(0.00)	
or more	Self	83.7	41.1	-0.37	-0.14	4.8	-4.1	0.99	0.12	3525
	Employed	(71.0)	(0.49)	(0.02)	(0.02)	(6.3)	(5.9)	(0.37)	(0.00)	
	Wage and	44.3	21.3	-0.40	-0.16	-1.1	-0.4	2.07	0.13	57172
All educ.	Salary	(38.0)	(0.48)	(0.00)	(0.00)	(0.6)	(0.6)	(0.13)	(0.00)	
groups	Self	68.3	34.2	-0.40	-0.15	-1.1	-1.6	0.12	0.13	6593
	Employed	(56.8)	(0.50)	(0.01)	(0.01)	(3.6)	(3.4)	(0.89)	(0.00)	

	Y^i	$\text{std}(\tilde{y}_1^i)$	β_y^i	$\text{corr}(\tilde{y}_1^i, \tilde{R})$
WS	80	8	13.33	0.25
SE(1)	80	8	26.67	0.50
SE(2)	80	8	-32.00	-0.60

Table 5: Characteristics of investors

	Y^h	$\text{var}(\tilde{y}_1^h)$	$\text{corr}(\tilde{y}_1^h, \tilde{R})$	β_y^h	C_{mv}^i	V_{mv}^i
SE	80	256	0.80	85.33	73	92
WS	80	225	0.00	0.00	80	225
A^*= 0.103		$RRA^* \approx 4$				

Table 6: Deriving the minimum variance portfolios

Table 7: This table shows calculations of the minimum variance consumption level ($C^i_{mv,t}$); the associated variance ($V^i_{mv,t}$); the degree of relative risk aversion (RRA) that would be required to make an individual indifferent between wage-earning and self-employment; and the proportion of the consumption gap between self-employment and wage earning "explained" by the model (ΔC). Calculations are based on the specification in Table 2, that is, the "five-year rule". All dollar magnitudes are measured in thousands of 1996 dollars.

SE status	Education	Y^i_t	$B^i_{y,t}$	$C^i_{mv,t}$	Ψ^i_t	$V^i_{mv,t}$	RRA	ΔC
All	All	780.2	-19.4	**781.2**	14.8	**186.5**	n.a.	n.a.
No more than	Wage and Salary	637.1	-246.4	**649.4**	15.1	**49.6**		
high school	Self Employed	954.3	-476.5	**978.2**	13.9	**247.4**	**127.86**	**1.17**
Some college	Wage and Salary	889.6	321.3	**873.6**	14.5	**104.3**		
or more	Self Employed	1418.1	-1470.5	**1491.6**	15.1	**1627.3**	**54.57**	**2.75**
All educ.	Wage and Salary	734.4	-112.7	**740.0**	14.8	**66.9**		
groups	Self Employed	1179.6	-886.0	**1223.9**	14.9	**903.4**	**60.59**	**2.48**

Table 8: For each specification, the figures in the column labeled RRA show the degree of relative risk aversion that would make an individual indifferent between self-employment and wage-earning. The figures in the column labeled (ΔC) show the proportion of the difference between consumption in self-employment and wage-earning that is “explained” by the model. The first set of numbers is based on the point estimates of β_y^i from Tables 2 and 4. In the second set of numbers, β_y^i is set equal to zero.

		No more than high school		Some college or more		All educ groups	
		RRA	(ΔC)	RRA	(ΔC)	RRA	(ΔC)
		Point Estimate					
5 year rule	Mean	127.9	(1.2)	54.6	(2.7)	60.6	(2.5)
	Median	105.2	(1.4)	43.0	(3.5)	50.0	(3.0)
Pseudo-IV	Mean	196.8	(0.8)	82.0	(1.8)	98.1	(1.5)
	Median	178.1	(0.8)	5.3	(28.4)	66.0	(2.3)
Irrevocable	Mean	61.8	(2.4)	49.7	(3.0)	42.7	(3.5)
	Median	34.4	(4.4)	47.4	(3.2)	34.9	(4.3)
		$\beta_y^i = 0$					
5 year rule	Mean	110.8	(1.4)	23.3	(6.4)	41.9	(3.6)
	Median	88.2	(1.7)	11.7	(12.8)	31.4	(4.8)
Pseudo-IV	Mean	97.7	(1.5)	126.2	(1.2)	86.2	(1.7)
	Median	79.1	(1.9)	49.2	(3.0)	54.1	(2.8)
Irrevocable	Mean	63.0	(2.4)	12.6	(11.9)	27.2	(5.5)
	Median	35.6	(4.2)	10.3	(14.6)	19.4	(7.7)

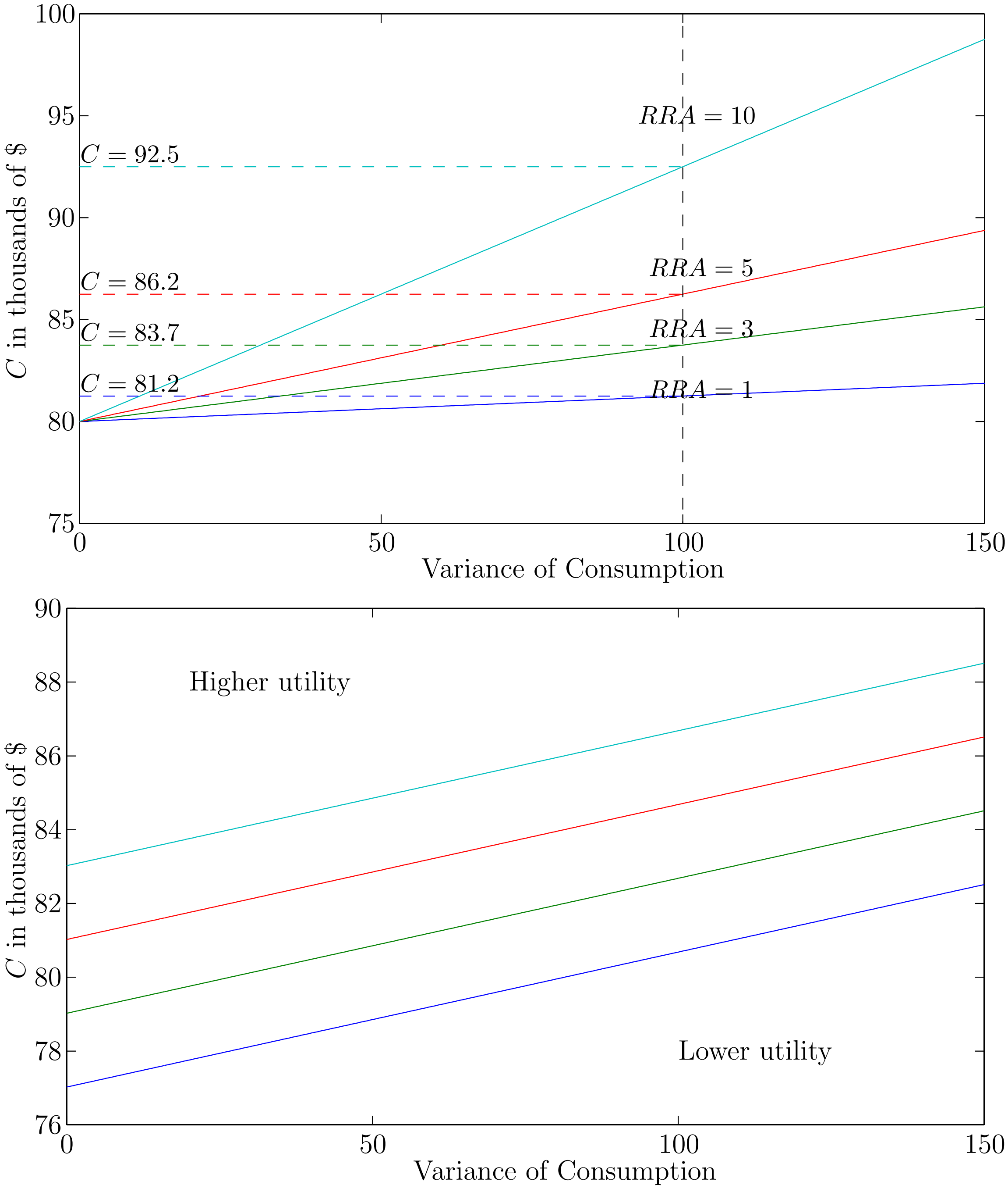

Figure 1: The upper panel shows indifference curves of consumers with different levels of absolute risk aversion. Higher ARA leads to steeper indifference curves. "RRA" is the approximate level of local relative risk aversion for a consumer with $\mathrm{E}(\tilde{c}_1) = 40$. The lower panel shows the indifference map of a single consumer with constant absolute risk aversion. ARA is chosen so that RRA is approximately 3.

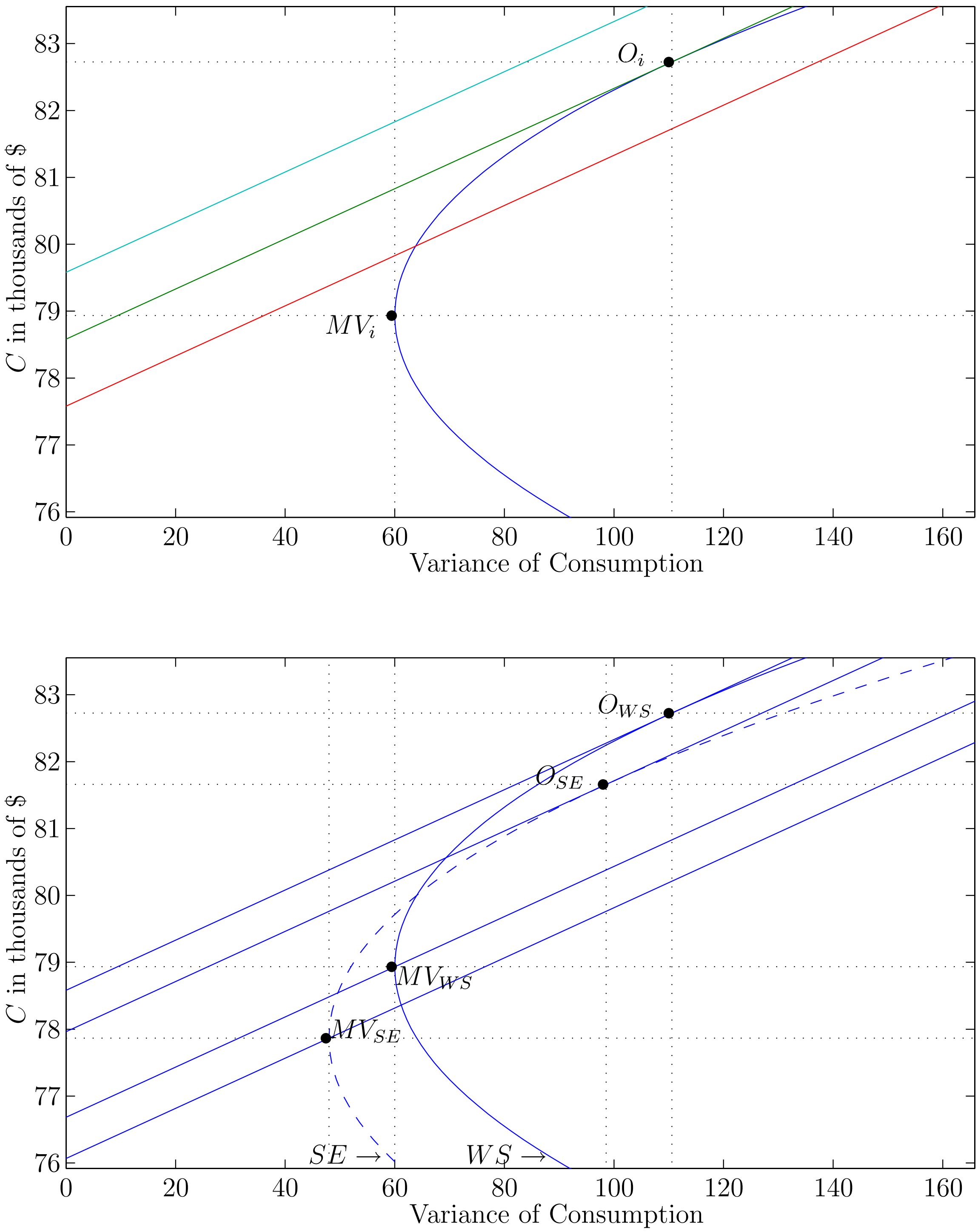

Figure 2: The top panel shows the portfolio choice decision, (O_i) for an individual who has chosen occupation i with income and preference properties listed in row "WS" of Table 5. The lower panel considers the same individual now faced with two different occupations with characteristics listed in rows "WS" and "SE(1)" of Table 5. WS and SE differ only with respect to the covariance of income and asset returns – $\text{corr}(\tilde{y}_1^{WS}, \tilde{R}) = .25$ and $\text{corr}(\tilde{y}_1^{SE}, \tilde{R}) = .5$. The individual prefers wage and salary employment over self-employment because the optimal point on the WS feasible set (O_{WS}) is on a higher indifference curve than the optimal point on the SE feasible set (O_{SE}).

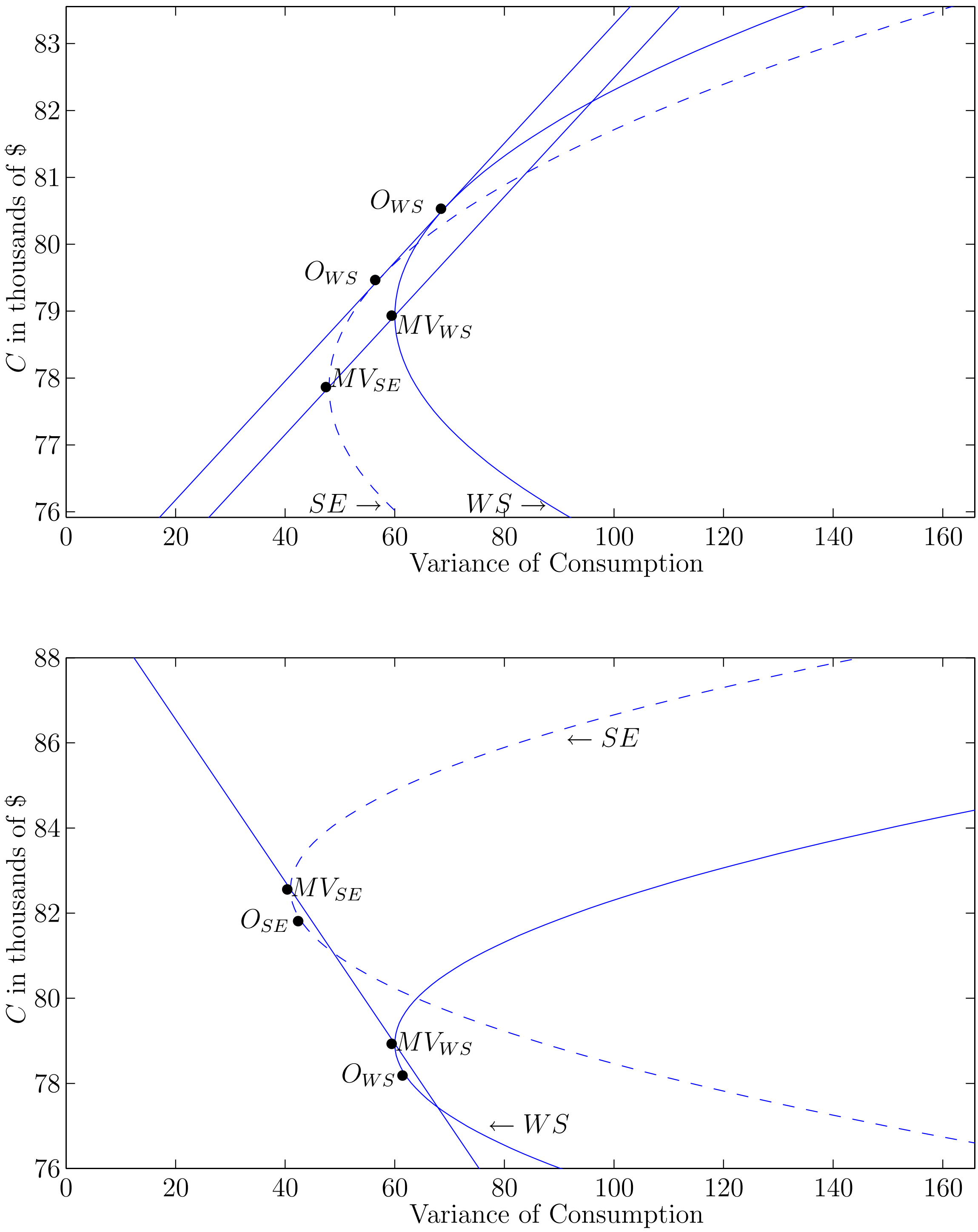

Figure 3: The top panel shows the indifference curves of an individual who is indifferent between self-employment and wage-and-salary employment (i.e. $A = A^*$ as described on page 16 of the text). Occupations WS and SE are described in rows "WS" and "SE(1)" of Table 5. The top panel also illustrates that if an investor is indifferent between the optimal points in two feasible sets, she is also indifferent between the minimum variance points. The bottom panel shows an example that compares occupation "WS" (again see row "WS" in Table 5) and occupation "SE" (now look at occupation "SE(2)" in Table 5). WS and SE again differ only with respect to the covariance of income and asset returns – $\text{corr}(\tilde{y}_1^{WS}, \tilde{R}) = .25$ and $\text{corr}(\tilde{y}_1^{SE}, \tilde{R}) = -.6$. For these parameter values, self-employment strictly dominates wage and salary employment because its lifetime variance is lower and its consumption level is higher. Only a risk-loving investor (with negative risk aversion) would choose wage-and-salary employment, as illustrated by the fact that the indifference curve through the minimum variance points for the two feasible sets has a negative slope.

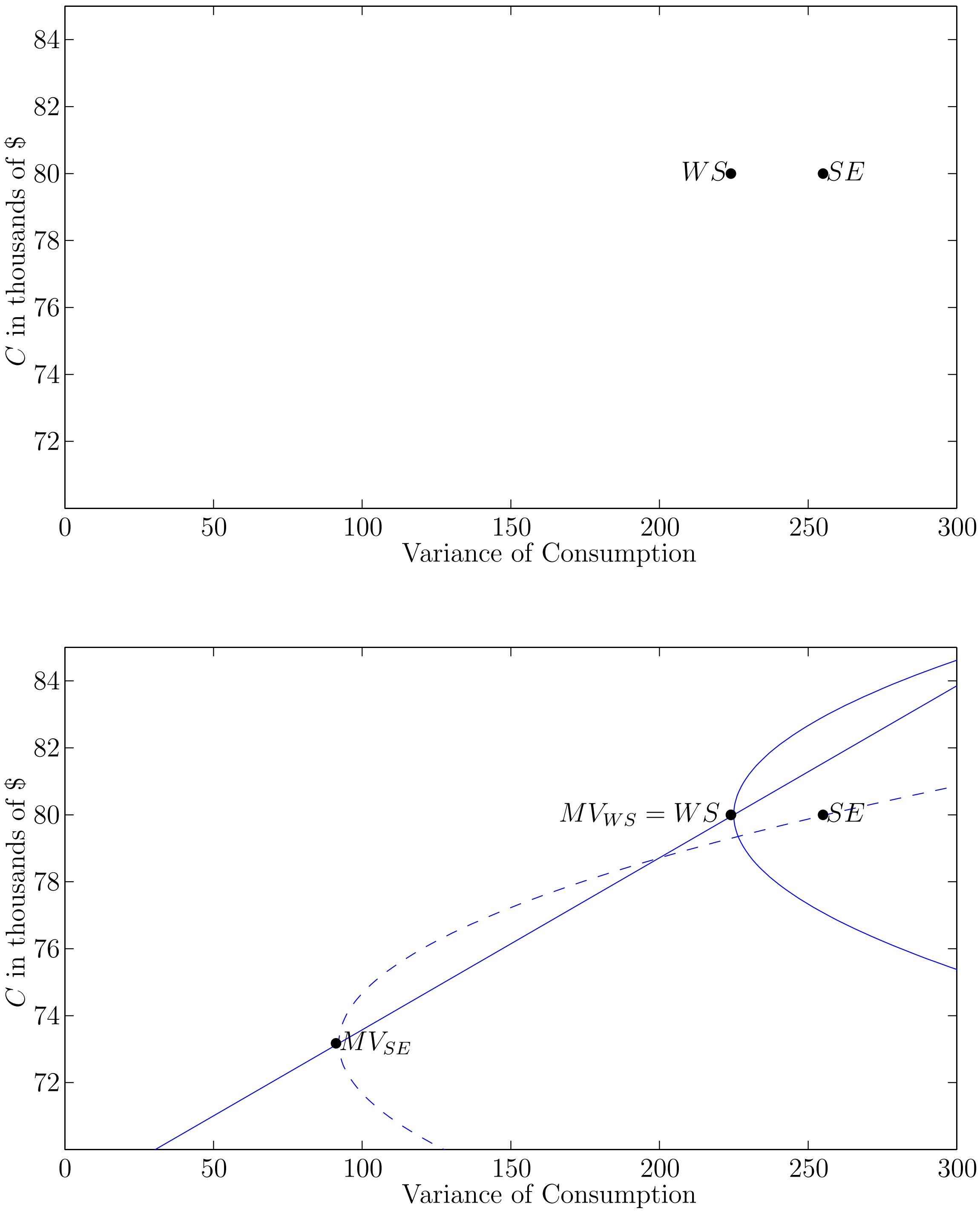

Figure 4: The top figure shows an example in which wage and salary employment has the same consumption but a lower variance than self-employment. Nevertheless, the bottom figure demonstrates that self-employment might be preferred because its minimum variance is lower than the minimum variance for wage and salary earning. Thus, for some preferences, self-employment will be preferred.