Identifying idiosyncratic career taste and skill with income risk

STEPHEN H. SHORE
Department of Economics
Johns Hopkins University
Baltimore, MD 21218
shore@jhu.edu

DANIEL BARTH
Department of Economics
Johns Hopkins University
Baltimore, MD 21218
dbarth4@jhu.edu

SHANE T. JENSEN
Department of Statistics
The Wharton School, University of Pennsylvania
Philadelphia, PA 19104
stjensen@wharton.upenn.edu

July 15th, 2010

Abstract

How important to well-being is choosing a career with the right fit? This question is difficult to answer because we observe individuals only in their chosen careers, but not in the other (presumably inferior) options they did not choose. To overcome this problem, we estimate a model in which individuals vary in risk tolerance, ability and idiosyncratic skill in and taste for various careers. Individuals choose from a set of careers that differ in income risk, typical pay, and other attributes. Given the model, the importance of idiosyncratic taste and skill is identified from the shift in the distribution of income risk with risk aversion. We estimate the model using individual-specific measures of income volatility to proxy for income risk and survey questions about hypothetical income gambles to proxy for risk preference, both from the PSID. We separate idiosyncratic career taste from skill using the pay gap between high- and low-income risk people with high and low risk-aversion.

---

1We thank Ellie Pavlick for excellent research assistance. We thank seminar participants at Dartmouth College, Brown University, the University of Maryland – College Park, Kyoto University, the Bureau of Labor Statistics, the Risk Theory Society 2010 meeting, and the SOLE/EALE 2010 meeting. Previous versions of this draft have circulated as “The Human Capital Risk-Return Menu” and “A Structural Model of Career Choice.”

2JEL: J24 (Human Capital; Skills; Occupational Choice; Labor Productivity)

3keywords: occupational choice, career choice, income risk, idiosyncratic taste and skill
1 Introduction and Motivation

The Road Not Taken

Two roads diverged in a wood, and I–
I took the one less traveled by,
And that has made all the difference.
–Frost (1920)

This paper aims to test the claim implicit in Frost’s poem.1 If we think of Frost’s “roads” as career paths, how important is road choice to well-being? How important is a chosen road’s idiosyncratic fit, of finding the road best suited to idiosyncratic tastes and skills? In the language of economists, how much would money would an individual demand to accept another road? We use a model of road choice to estimate the distribution of road quality.

Such a model for career choice is challenging for three reasons. First and most obviously, we never observe the full choice set; we observe only the career that was choosen, but not (presumably inferior) options that weren’t chosen.2 Second, data on the the chosen career’s attributes may be limited. Third, career choice decisions depend substantively on non-pecuniary factors (e.g. working conditions, interest in the work, etc.) that economists typically don’t observe and about which we have few interesting, quantitative hypotheses.

We overcome these problems using variables we do observe and about which we have well-developed models: a chosen career’s income risk and a person’s coefficient of relative risk aversion. While risk-related variables may affect career choice, they are hardly the only drivers of such choice. We use them to infer the

1In fact, a large body of poetry criticism argues that Frost’s intended meaning was not the literal and commonly believed one (Pritchard, 1984). Scholars note “Frost’s decision to make his two roads not very much different from one another, for passing over one of them had the effect of wearing them ‘really about the same.’” (Monteiro, 1988)
2This is merely a career choice application of the classic problem of measuring treatment effects (surveyed in Holland, 1986).
importance of other factors, including the unobservable ones that we model as idiosyncratic. This is possible if we assume that risk aversion is an attribute of a person (and therefore not affected by the career they choose) and income risk is an attribute of a career (and therefore not affected by the risk aversion of the person who chooses it). Violation of these identifying assumptions in the directions one might expect will bias our estimates downward, so that our quite substantial estimates may be viewed as a lower bound.

We envision a model in which people are endowed with a preference for risk and level of overall ability, broadly conceived as risk-free, career-independent earning potential. Careers differ in income risk and typical, worker-independent pay. We allow individuals to have idiosyncratic taste for – or skill and correspondingly higher pay in – some careers over others. Income risk is resolved after the career is chosen. Both careers and individuals may differ in other attributes, which may be unobservable to the econometrician though all are observable to workers; we must place strong restrictions on the relationship between these unobservables and observed income risk and risk aversion. We abstract from search frictions and bounded rationality concerns; workers observe and understand the full set of career options from which they can choose. We also abstract from intertemporal concerns; workers chose careers only once.

*Ceteris paribus,* the relationship between income risk and risk-aversion will be weakly negatively monotonic; risk tolerant individuals will choose the riskiest careers (which will carry a compensating wage differential for income risk) while risk intolerant individuals will choose the safest careers (as shown in Figure 1).³ The local risk-return trade-off (the marginal risk premium) is determined by the risk-aversion (slope of risk-return indifference curve) of the marginal individual at that quantity of risk. There is already a large empirical and theoretical litera-

³Deleire and Levy (2004) present a similar figure in which people with heterogeneous preference for injury (not income risk) sort into safer and riskier (by probability of injury) jobs.
The solid line represents a menu of risk-return career options for a set of careers that are otherwise identical. The dotted line presents a risk-return indifference curve for a hypothetical individual with a high risk tolerance; the dashed line presents a risk-return indifference curve for a hypothetical individual with a low risk tolerance.

While we estimate this relationship in Table 6 for our data, this is not our primary goal. Instead, we aim to see if this well-studied and well-understood risk-return relationship can be used to identify other parameters in a career choice model.

Individuals have idiosyncratic taste for some careers over others and idiosyncratic taste on this risk-return relationship. While we estimate this relationship in Table 6 for our data, this is not our primary goal. Instead, we aim to see if this well-studied and well-understood risk-return relationship can be used to identify other parameters in a career choice model.

---

idiosyncratic aptitude (and therefore income) in some careers over others. This idiosyncratic taste or skill will lead some highly risk-averse people to choose careers with high income risk (that they happen to love or at which they excel); similarly, it will lead some risk-tolerant people to choose careers with low income risk. The insight of this paper is that the risk-return model in the previous paragraph allows us to identify the dispersion of these idiosyncratic factors. We know the welfare cost of deviating from an anticipated income risk choice for someone with a given risk preference; by observing the distribution of such deviations we can back out the dollar-equivalent value of the idiosyncratic tastes or skills that made these deviations optimal. In particular, the dispersion of (perceived) idiosyncratic skill and taste is identified from shifts in the distribution of income risk as risk-aversion changes. The larger the drop in income risk (relative to the conditional distribution of income risk) as risk aversion rises, the lower the implied dispersion of idiosyncratic values. We estimate a lower bound of 45% of income on the standard deviation of these values.

This framework provides an application for the recent literature on heterogeneity in income volatility. Income volatility is frequently used as a proxy for income risk. Meghir and Pistaferri (2004) and Alvarez, Browning, and Ejrnaes (2001) show that income volatility differs across individuals. Jensen and Shore (2009a,b) estimate the distribution of ex-ante, individual-specific volatilities in the population. The 1996 PSID includes a measure of self-reported risk tolerance, elicited from a survey asking the individual if they would take a series of large gambles (Barsky, Juster, Kimball, and Shapiro, 1997; Sahm, 2007; Kimball, Sahm, and Shapiro, 2008, 2009). We merge these risk-aversion values with estimates of individuals’ volatilities in the PSID, both from Jensen and Shore (2009a,b) and also from Meghir and Pistaferri (2004). Individuals who self-identify as risk tolerant tend to have more volatile income streams. At the same time, we observe a non-degenerate joint
distribution of income volatility and risk tolerance; conditional on observed risk
tolerance, individuals choose a wide variety of levels of income volatility.

Our parametric and statistical assumptions are designed to imply a logit struc-
ture (McFadden, 1974). Logit models have long been used in the reduced-form
occupational choice literature to study the relative importance of covariates on
choice from a finite list of observed careers (Boskin, 2004; Field, 2009). Without
an economic model, the multinomial logit setting can be identified only up to a
normalization: doubling the utility from all careers (doubling all coefficients and
the error term) has no affect on career choice. Our model provides a normaliza-
tion with an economic meaning so that all estimates can be expressed in terms of
their (log) certainty equivalents. Because we assume a continuum of careers,
our model has the continuous logit structure previously used to study home lo-
cation choice (Ben-Akiva, Litinas, and Tsunokawa, 1985). In the housing applica-
tion, there is effectively a continuum of homes located on a two-dimensional plane;
while researchers may observe few home attributes, they do observe the home’s
location in this plane. In our setting, we don’t observe the exact chosen career, but
we observe where that career is located along an income risk “line.”

Keane and Wolpin (1997) provides an alternative – and very different – normalizing model of
optimal educational investment and subsequent choice from five broad career categories.
The common reduced-form alternative is to use pay as one of the covariates that influences
occupational choice, giving other coefficients a dollar-equivalent interpretation (Willis and Rosen,
1979; Robertson and Symons, 1990; Siow, 1984). This alternative assumes away the problem that
we do not observe occupational pay, but rather the pay of those who choose an occupation. High
pay in a given occupation may reflect not just high pay for that occupation but also high ability (or
idiosyncratic individual-career-specific skill) among those who choose that occupation. In some
sense, this assumption is the first-moment analog to the one we make in this paper; we assume that
income risk (the second moment of pay) is associated with a career and not the person that chooses
this career; to use pay as a numeraire, we must assume that expected pay (the first moment of pay)
is associated with a career and not the person who chooses that career.

Dynamic data can be used to overcome the problem that ability may be correlated with pay; pay
changes resulting from occupational changes can be used to estimate career-specific effects holding
overall individual ability fixed (Stinebrickner, 2001). This approach does not tackle the problem
that idiosyncratic individual-career-specific skill may change at such transitions, with some careers
systematically receiving workers for whom that career is a better “fit”.

While most papers on occupational choice rely on choice from a finite and specified list of
occupational options, some papers allow the number of occupations to be very large, using occu-
pational categories primarily to identify the attributes of chosen careers. For example, Deleire and
Naturally, our economic normalization of a logit model inherits the econometric limitations of any logit model; anything that will bias logit coefficient estimates will be a problem for us. To restrict career-specific unobservables, we must assume that career-specific unobservables that are not independent of income risk have the same affect on career choice regardless of risk aversion. To restrict individual-specific unobservables, we must assume that individual-specific unobservables that affect a career’s value to an individual do not vary with the income risk of that career; this rules out the possibility that the risk distribution of career options varies across individuals. In the conclusion, we discuss how our results should be interpreted if this assumption is violated. Most significantly, we require that risk aversion is associated with a person, and income risk is associated with a career. In truth, income risk may depend on individual attributes as well, and might be correlated with risk-aversion. In this case, to the degree that risk-averse people will make any career less risky, our estimates of the variance of idiosyncratic factors will be biased downward and can be viewed as a lower bound. Similarly, we might imagine that risky careers lead individuals to become more risk tolerant; again, this biases estimates of the importance of idiosyncratic factors downward.

The model allows us to use the joint distribution of income risk and risk-aversion – in particular, the shift in the conditional distribution of income risk as risk-aversion changes – to estimate the variance of idiosyncratic taste and skill together. We can separate taste from skill using income data. When a risk-averse person chooses a career with substantial income risk, on average he must be compensated in some way for this risk. Such compensation could be in the form of higher idiosyncratic skill in this career (and therefore higher pay) or higher id- 

---

Levy (2004) use 46 occupational codes to map their occupational attribute of interest (injury and fatality risk) to individuals who choose those occupations. In our case, panel data allow us to obtain individual-specific estimates of our career attribute of interest (income volatility as a measure of income risk), so we can examine career attribute choice without explicitly observing the chosen career.
iosyncratic taste for this career (and therefore higher enjoyment). To the degree that idiosyncratic productivity shocks are larger than idiosyncratic taste shocks, we should see risk-averse people with high income risk earning more than risk-averse people with low income risk. By comparing this high-income risk versus low-income risk pay gap for those with high and low risk-aversion, we can difference out market-wide compensating differentials for income risk. Since we observe a similar pay gap for those with high and low risk-aversion, we infer that nearly all idiosyncratic variation is in career taste and not career skill.

2 Model

We present a model of career choice over risky careers. Individuals choose from a set of career options. Each career option has a quantity of income risk, a typical pay for that career, and other non-pecuniary attributes. Each individual has a preference for income risk, an overall ability (typical pay for that individual), and other attributes. There is a distribution of career options and a distribution of people in the population. In addition to these innate traits of careers and individuals, there are traits specific to an individual in a given career. Some individuals have an idiosyncratic taste for some careers over others; also, some individuals are idiosyncratically better (higher productivity, and therefore higher pay) in some careers than others. From the set of career options, each individual makes a one-time, irrevocable choice of the best career. Then, career-specific income risk is realized. In our data, we observe proxies for individuals’ risk preferences and their chosen careers’ income risks.
2.1 Setup

2.1.1 Careers

Career options are indexed by \( c \in \{1, \ldots, N_C\} \). Careers have four attributes, \( X^C \equiv \{\sigma^2_c, y^C_c, x^{CO}_c, x^{CU}_c\} \); \( X^C_c \equiv \{\sigma^2_c, y^C_c, x^{CO}_c, x^{CU}_c\} \) is the set of attributes for career \( c \). \( \sigma^2_c \) is a measure of the income risk in career \( c \). \( y^C_c \) is a career-specific measure of log pay in career \( c \). \( x^{CO}_c \equiv [x^{CO}_c; x^{CU}_c] \) is a vector of covariates or attributes of career \( c \); \( x^{CO}_c \) are the attributes observable to the econometrician and to workers; \( x^{CU}_c \) are the set of attributes observable to workers but not to the econometrician. \( N_{CO} \) and \( N_{CU} \) are the number of observable and unobservable attributes, respectively.

Later, we will consider a continuum of atomistic careers, so that \( N_C \to \infty \). In this case, \( f^C(X^C) \) is the distribution of career attributes, taken over the set of possible careers. Naturally, in equilibrium some careers will be chosen more than others, so that \( f^C \) will typically not be the distribution of the attributes of chosen careers; the distribution of career options and the distribution of career choices need not be the same.

2.1.2 People

People are indexed by \( i \in \{1, \ldots, N_I\} \). People have four attributes, \( X^I \equiv \{\gamma_i, y^I_i, x^{IO}_i; x^{IU}_i\} \); \( X^I_i \equiv \{\gamma_i, y^I_i, x^{IO}_i; x^{IU}_i\} \) is the set of attributes for person \( i \). \( \gamma_i \) is a measure of risk-aversion for person \( i \). \( y^I_i \) is a person-specific measure of log pay (general ability or productivity) for person \( i \). \( x^I_i \equiv [x^{IO}_i; x^{IU}_i] \) is a vector of covariate attributes of person \( i \); \( x^{IO}_i \) are the set of attributes observable both to the econometrician and to workers in the model; \( x^{IU}_i \) are the set of attributes observable to workers in the model but not to the econometrician. \( N_{IO} \) and \( N_{IU} \) are the number of observable

---

8 The assumption that income risk is associated with a career, not an individual, is a strong one. Jacobs, Hartog, and Vijverberg (2009) discusses the biases associated with making this assumption in reduced-form risk-return estimation.
and unobservable attributes, respectively. \( f^I(X^I) \) is the distribution of peoples' attributes in the population.

### 2.1.3 Individual-Career-Specific Fit

We assume that some careers are a better fit for some people than others. Fit is characterized by two attributes, \( X \equiv \{ y, l \} \). \( X_{i,c} \equiv \{ y_{i,c}, l_{i,c} \} \) is the fit for person \( i \) in career \( c \). \( y_{i,c} \) is an individual-career-specific measure of log pay (idiosyncratic productivity) of person \( i \) in career \( c \). \( l_{i,c} \) is an individual-career-specific measure of idiosyncratic enjoyment of person \( i \) in career \( c \).

\( f_{i,c}(X) \) is the joint distribution of \( X_{i,c} \equiv \{ y_{i,c}, l_{i,c} \} \). We require that \( X_{i,c} \) and \( X_{i,c'} \) be identically distributed and independent of one another when \( c \neq c' \), and also that \( X_{i,c} \) be independent of \( X^I_i \) and \( X^C_c \). Independence when \( c \neq c' \) is the standard "independence of irrelevant alternatives" assumption present in multinomial logit settings. Independence across \( i \) is also required for inference when we estimate the model on data.

We place very limited restrictions on the distribution of \( X_{i,c} \). We do not require that \( y_{i,c} \) and \( l_{i,c} \) be independent of one another. We merely require that the cdf of \( y_{i,c} + \frac{\alpha}{1-\alpha} l_{i,c} \) (with a particular \( \alpha \in [0,1] \) we define in Section 2.1.4) be twice differentiable (de Haan and Ferreira, 2006). The normal and exponential distributions are examples of such distributions. Coupled with the independence assumptions from the previous paragraph, this implies that \( \lim_{N \to \infty} \max_{c \in C} (y_{i,c} + \frac{\alpha}{1-\alpha} l_{i,c}) \) has an extreme value distribution (of Type I, Gumbel). Alternatively, we can just assume that \( y_{i,c} + \frac{\alpha}{1-\alpha} l_{i,c} \) has an extreme value distribution. In either case, our aim is to estimate the scale parameter \( \beta \) that governs the distribution either of an extreme-valued \( y_{i,c} + \frac{\alpha}{1-\alpha} l_{i,c} \) or the extreme-valued maxima. When \( y_{i,c} + \frac{\alpha}{1-\alpha} l_{i,c} \) has an extreme value distribution \( \text{var}(y_{i,c} + \frac{\alpha}{1-\alpha} l_{i,c}) = \beta^2 \pi^2/6; \) when \( y_{i,c} + \frac{\alpha}{1-\alpha} l_{i,c} \) does not have an extreme value distribution but its maxima does, \( \text{var}(y_{i,c} + \frac{\alpha}{1-\alpha} l_{i,c}) \propto \beta^2 \pi^2/6. \)
2.1.4 Preferences

The model that follows assumes risk-averse, expected utility maximizing individuals who care about *stochastic* income $Y$ and career enjoyment $L$ (for leisure). Individuals have Cobb-Douglas preferences over $Y$ and $L$, and expected utility preferences over the composite, $v$. Individual $i$ in career $c$ has an expected utility of:

$$ Eu(i, c) = E \left[ \frac{v_{i,c}^{1-\gamma_i}}{1-\gamma_i} \right] $$

(1)

$$ v_{i,c} \equiv Y_{i,c}^{1-\alpha} L_{i,c}^\alpha $$

(2)

$$ \ln Y_{i,c} \equiv y_i^C + y_i^l + y^x(x_i^l, x_c^C) + y_{i,c}^\varepsilon + \sigma_c \xi - \frac{1}{2} \sigma_c^2 $$

(3)

$$ \ln L_{i,c} \equiv l^x(x_i^l, x_c^C) + l_{i,c}^\varepsilon $$

(4)

Composite felicity (equation 2) $(v)$ is a Cobb-Douglas function of income $Y_{i,c}$ and career enjoyment $L_{i,c}$. The relative importance of income and career enjoyment is determined by $\alpha$. Note that we impose an elasticity of substitution of one and do not allow heterogeneity in $\alpha$.

For simplicity, we assume a one-period model in which income $Y$ is merely equal to consumption. Log income in equation 3 is the sum of: career-specific pay $(y_i^C)$, including a premium for size, risk, or non-pecuniary attributes; individual-specific pay or overall productivity $(y_i^l)$; the affect of the interaction of individual- and career-specific covariates on pay $(y^x(x_i^l, x_c^C))$; individual-career-specific pay $(y_{i,c}^\varepsilon)$, the individual’s career-specific productivity; and, the realization of a shock $(\xi)$ with a standard normal distribution, adjusted so its exponentiated expectation is one. Log enjoyment is the sum of: the affect of the interaction of individual- and

---

9It is straightforward extend this to a multi-period setting in which a one-time career decision affects income dynamics, and consumption and saving respond optimally to income shocks. This richer structure loses the clean analytic framework presented below, but it is easy to implement numerically. We omit it here for parsimony.
career-specific covariates on enjoyment \((l^x(x^I_i, x^C_c))\); and, individual-career-specific enjoyment \((l^\epsilon)\). Covariates affect the pay and enjoyment of individual \(i\) in career \(c\) as follows:

\[
y^x(x^I_i, x^C_c) \equiv l'(\theta^y \cdot (x^I_i x^C_c)) \; \nu; \; l^\epsilon(x^I_i, x^C_c) \equiv l'(\theta^l \cdot (x^I_i x^C_c)) \; \nu^{10}
\]

2.1.5 Career Value

Plugging equations (2), (3), and (4) into equation (1), evaluating the expectation, and transforming yields a log income certainty equivalent measure of the value of career \(c\) to person \(i\):

\[
V(i, c) \equiv \ln \left( \frac{(1 - \gamma_i) \; Eu}{(1 - \alpha)(1 - \gamma_i)} \right) = y^I_i + y^C_c + y^x(x^I_i, x^C_c) + \varepsilon_{y,i,c}
\]

\[
+ \frac{\alpha}{1 - \alpha} \left( l^\epsilon_{i,c} + l^x(x^I_i, x^C_c) \right) - \frac{1}{2} (\alpha + \gamma_i - \alpha \gamma_i) \sigma_c^2
\]

Individuals will choose the career with the highest \(V\). Note that the problem is set up so that person-specific ability \((y^I)\) has no impact on the career chosen; it merely shifts the value of all careers equally. In our data, we observe the curvature parameter of the utility function with respect to income or consumption. And while the importance of idiosyncratic career preference is apparent in the data, we cannot separate large idiosyncratic career taste shocks from important

\[10\text{The affect of covariates on pay or enjoyment depends on coefficient matrices, } \theta^y \text{ and } \theta^l:\]

\[
\theta^y = \begin{bmatrix} \theta^y_{OO} & \theta^y_{OU} \\ \theta^y_{UO} & \theta^y_{UU} \end{bmatrix}; \; \theta^l = \begin{bmatrix} \theta^l_{OO} & \theta^l_{OU} \\ \theta^l_{UO} & \theta^l_{UU} \end{bmatrix}
\]

\(\theta^y\) is a matrix of size \(NIO + NUC\) by \(NCIO + NCUC\). To explain the notation here, \(\theta^y_{OU}\) is a \(NIO\) by \(NCU\) matrix that gives the impact on pay of the interaction of observable individual attributes with unobservable career attributes. Note that “\(\cdot\)” indicates pair-wise multiplication, and \(\nu\) is a vector of ones so that we merely sum up all elements of the matrix of affects, \(\theta^y \cdot (x^I_i x^C_c)\) or \(\theta^l \cdot (x^I_i x^C_c)\).
idiosyncratic career taste shocks. This informs the following transformations:

\[ \tilde{\gamma}_i \equiv \alpha + \gamma_i - \alpha \gamma_i \]  

(8)

\[ l^x(x^I_i, x^C) \equiv \frac{\alpha}{1 - \alpha} l^x(x^I_i, x^C); \tilde{l}_{i,c} \equiv \frac{\alpha}{1 - \alpha} \tilde{l}_{i,c} \]  

(9)

We can then re-write the value of each career as:

\[ V(i, c) = y^I_i + y^C_c + \gamma^x(x^I_i, x^C_c) + l^x(x^I_i, x^C_c) - \frac{1}{2} \tilde{\gamma}_i \sigma^2_c + y^\varepsilon_{i,c} + \tilde{l}^\varepsilon_{i,c} \]  

(10)

If we group pecuniary and non-pecuniary idiosyncratic terms and also group pecuniary and non-pecuniary covariate terms as

\[ \varepsilon_{i,c} \equiv y^\varepsilon_{i,c} + \tilde{l}^\varepsilon_{i,c} \text { and } v(x^I_i, x^C_c) \equiv y^v(x^I_i, x^C_c) + l^v(x^I_i, x^C_c) \]  

(11)

then

\[ V(i, c) = y^I_i + y^C_c + v(x^I_i, x^C_c) - \frac{1}{2} \tilde{\gamma}_i \sigma^2_c + \varepsilon_{i,c}. \]  

(12)

Equation (12) gives career choice a standard, random utility, multinomial logit structure (McFadden, 1974). Individuals choose the career that gives them the highest utility, and this utility depends on career attributes that affect everyone equally (here, \( y^C \)), career attributes that affect different individuals differently (here, observable and unobservable covariates \( v(x^I_i, x^C_c) \) and the utility cost of income risk \( \sigma^2_c \) that depends on risk aversion \( \tilde{\gamma}_i \)), and an error term (here, \( \varepsilon_{i,c} \equiv y^\varepsilon_{i,c} + \tilde{l}^\varepsilon_{i,c} \)).

What is unique here is that our economic model provides an economic interpretation to the utility level normalization (log of certainty equivalent income) and the coefficient on the \( \sigma^2_c \times \tilde{\gamma}_i \) term (\( -\frac{1}{2} \)). Because the economic model provides our normalization, coefficient estimates and the variance of the error term now have
an absolute, log-income-equivalent meaning.

2.2 Stylized Model Without Idiosyncratic Career Preference

We begin by considering a model without idiosyncratic career taste or covariates, so that \( y_{i,c}^e = l_{i,c}^e = 0 \). In this setting, there is no person-specific heterogeneity conditional on \( \gamma \) (save for heterogeneity in \( y^I \) which does not affect choice). All individuals with the same \( \gamma \) are indifferent among any choices they choose with positive probability.

This implies a weakly (negatively) monotonic relationship between risk-aversion and income risk choice. We should never see a more risk tolerant person choosing less income risk. Here, we consider a continuum of careers on some range of \( \sigma^2 \), which have full support in the sense that all careers are chosen by someone. Let \( \gamma(\sigma^2) \) be the risk-aversion of the person who chooses income risk \( \sigma^2 \).

At an interior optimum, the individual’s first order condition (holding career enjoyment fixed) requires that\(^{12}\)

\[
\frac{dy^C}{d\sigma^2} = \frac{1}{2} \tilde{\gamma}
\]

Equation (13) must hold for each \( \{\sigma^2, \tilde{\gamma}\} \) pair and we know the risk aversion of the marginal individual for each \( \sigma^2 \), then we can trace out \( y^C \) as

\[
y_c^C = y_0^C + \frac{1}{2} \int_0^{\sigma^2} \tilde{\gamma}(x)dx
\]

Here, \( y_0^C \) is the sum of log pay and enjoyment for a risk-free career. Note the

\(^{11}\)Because of the full support assumption, each \( \sigma^2 \) is chosen by someone and therefore maps to a \( \gamma \) though a measure zero set of \( \sigma^2 \) values may map to multiple \( \gamma \) values. The fact that the number of such points is of measure zero means that the values we use here do not affect the risk-return menu.

\(^{12}\)Obtained by differentiating expected utility in equation (12) with respect to \( \sigma^2 \), setting equal to zero, and rearranging terms.
strong assumptions needed here, namely that all individuals face the same risk-return menu (up to an ability intercept which can differ across individuals). A graphical depiction of this menu is given in Figure 1.

2.3 Incorporating Idiosyncratic Career Preference

The stylized model in Section 2.2 has a homogeneous risk-return menu. Consequently, we should never observe an individual with higher risk aversion choosing higher income risk. This is wildly at odds with the data, which shows substantial heterogeneity in the volatility observed by individuals with the same survey-based estimate of risk aversion. We model this heterogeneity in equation (12), where $X_{i,c}^{\varepsilon} \equiv \{y_{i,c}^{\varepsilon}, \tilde{l}_{i,c}^{\varepsilon}\}$ are idiosyncratic individual-career-specific productivity and taste shocks (for person $i$ in career $c$), respectively.

We require either that underlying idiosyncratic terms have an extreme value distribution, or that the number of careers $N_C$ be large in the sense that we can use extreme value theory to describe the best career, given the independence and distributional assumptions from Section 2.1.3. As a result, $\lim_{N_C \to \infty} \max_{c \in \{1, \ldots, N_C\}} \{\varepsilon_{i,c}\}$ has an extreme value (Type I) distribution.\(^\text{13}\)

Let $r$ refer to the set of careers in a rectangle $\{y^C + v(x_i^I, x^C), \sigma^2\}$ for person $i$. Let $s_r$ be the share of careers that fall in region $r$. We assume that the number of careers in each region $r$ is large enough that $\max_{c \in r} \varepsilon_{i,c}$ has an extreme value distribution (with scale parameter $\beta$). Consider the choice among careers $c$ in range $r$. Taking the size of the rectangle to zero, within-range differences between careers $c$ in $X^C$, will be trivially small. As a result, if the individual chooses a

\(^{13}\)There are two technical advantages to an extreme value approach. First, increasing the number of careers affects only the location parameter $\mu$, shifting the whole distribution up while leaving its shape (governed by parameter $\beta$) unchanged. As a result, we can normalize out $\mu$, so that we need not take a position on the total number of careers $N_C$ (an idea without precise meaning) to identify the model. Second, results are not dependent on a particular parametric shape for the distribution of individual-career-specific shocks, $\varepsilon_{i,c}$.
career from within range \( r \), it will be the one with the highest \( \varepsilon_{i,c} \); these within-range maxima will have an extreme value distribution with scale parameter \( \beta \). Alternatively, we can assume each individual \( \varepsilon_{i,c} \) has an extreme value distribution (again with scale parameter \( \beta \)), in which case we need not assume the number of careers in each range becomes large.

Given the extreme value distribution, the probability that an individual’s preferred career will lie in range \( r \) is

\[
\text{prob}(V(i, r) \geq V(i, q), \forall q \neq r) \propto s_re^{y^C + v(x^I, x^C) - \frac{1}{2} \tilde{\gamma}_i \sigma^2 / \beta}. \tag{15}
\]

The probability that a given range will have the highest value (equation 15) is nothing more than the pdf, the joint distribution of attributes \( X^C \) of careers chosen given \( X^I \). We re-write equation (15) by taking the size of each range to zero, so that the sums become integrals and \( s_r \) becomes \( f^C(X^C) \):

\[
f(X^C | X^I) \propto f(X^C | \tilde{\gamma} = 0, X^I) e^{-\frac{1}{2} \tilde{\gamma} \sigma^2 / \beta} \tag{17}
\]

\[
f(X^C | \tilde{\gamma} = 0, X^I) \propto f^C(X^C)e^{y^C + v(x^I, x^C)}/\beta \tag{18}
\]

The model implies that a risk-neutral person (equation 18) will choose careers proportional to their frequency \( f^C(X^C) \). *Ceteris paribus*, a risk-neutral person will be twice as likely to choose a career with a given set of attributes if twice as many careers have those attributes. A risk-neutral person is also more likely to choose careers with higher career-specific pay and enjoyment \((y^C + v(x^I, x^C))\). These

\[\mu \text{ and } \beta \text{ are the location and scale parameters of the extreme value distribution, and } \gamma_{em} \approx 0.577 \text{ refers to the Euler-Mascheroni constant.}\]
career-specific attributes dominates career frequency when idiosyncratic career preference is small ($\beta \to 0$). Without idiosyncratic preference, risk-neutral people merely choose the career with the highest $y^C + v(x^I, x^C)$; the distribution of risk choices is extremely tight around the “best” choice. However, as idiosyncratic career preference becomes larger ($\beta \to \infty$), careers are chosen only in proportion to their frequency; the distribution of choices becomes as diffuse as the distribution of careers $f^C$. We should be unsurprised to see that individual-specific ability ($y^I$) does not affect career choice as it increases the benefit of all careers equally.

Of course, we don’t observe all elements of $X^C$ or $X^I$. We begin by integrating out unobservable components of careers. To do so, we require that $\bar{\gamma}$ does not affect the expected payoff of some risk levels more than others, so that

$$E \left[ e^{(y^C + v(x^I, x^C))/\beta} \mid x^I, \sigma^2, x^{CO} \right]$$

(19)
does not vary with $\bar{\gamma}$. Then, we integrate out unobservable individual attributes. To do so, we must make the strong assumption that any individual unobservables must not vary with career risk, so that

$$E \left[ e^{(y^C + v(x^I, x^C))/\beta} \mid x^{IO}, x^{CO}, x^{CU} \right]$$

(20)
should not vary with $\sigma^2$. For example, unobservably high- and low-ability people must have career options with the same income risk distribution. In this case,

---

15We do not observe career-specific pay ($y^C$) (only total pay which includes individual-specific ability, $y^I$, and an idiosyncratic productivity shock to the chosen career, $y^c_{i,c}$) or unobservable career-specific attributes ($X^{CU}$). We integrate these out to obtain the marginal distribution of observable career choices.
equations (17) and (18) become:

$$f(\sigma^2 | \tilde{\gamma}, x^{IO}, x^{CO}) \propto f(\sigma^2 | \tilde{\gamma} = 0, x^{IO}, x^{CO})e^{-\frac{1}{2}\tilde{\gamma}\sigma^2/\beta}$$  \hspace{1cm} (21)

$$f(\sigma^2 | \tilde{\gamma} = 0, x^{IO}, x^{CO}) \propto f^C(\sigma^2 | x^{CO})E\left[e^{(y^C + v(x^I, x^C))/\beta} | x^{IO}, x^{CO}\right]$$  \hspace{1cm} (22)

The critical insight from equations (17) or (21) is that the distribution of risk choices for risk-averse people $f(\sigma^2 | \tilde{\gamma})$ is completely determined by the distribution for risk-neutral people $\tilde{\gamma} = 0$ and a single parameter $\beta$. Each conditional distribution $f(\sigma^2 | \tilde{\gamma})$ for a given $\tilde{\gamma}$ is merely an exponential shift of another such conditional distribution for another $\tilde{\gamma}$. The degree of that shift is governed by $\beta$, which is proportional to the standard deviation of the idiosyncratic individual-specific-career taste and skill shocks. As these shocks become larger, the shift becomes more modest and conditional distributions look more similar to one another (and more similar to the distribution of careers, $f^C$). As these shocks become smaller, the shift becomes more substantial and conditional distributions for high and low $\tilde{\gamma}$ become more different (and each becomes more concentrated around the “best” choice for that $\tilde{\gamma}$). When idiosyncratic shocks are large, the distribution of risk choices by risk-neutral people will reflect primarily the distribution of career options ($f^C(\sigma^2 | x^{CO})$); when idiosyncratic shocks are small, the distribution of risk choices by risk-neutral people will reflect primarily which risk values have careers with the highest expected (pecuniary and non-pecuniary) value.

Note that this model is highly overidentified when we observe the joint distribution of $\sigma^2$ and $\tilde{\gamma}$. One conditional distribution completely determines the shape of all the others, and a single parameter determines the link between all conditional distributions.
2.4 Idiosyncratic Career Taste vs. Idiosyncratic Career Skill

Equation (21) provides a way to estimate $\beta$ from the degree to which the conditional distribution of risk choices shifts with risk-aversion. $\beta$ measures the standard deviation of $y_{i,c} + \tilde{l}_{i,c}$. Without additional information, we cannot separate the relative importance of individual-specific shocks to skill in specific careers ($y_{i,c}$) from individual-specific shocks to taste for (enjoyment of) those careers $\tilde{l}_{i,c}$.

However, we can separate these two affects using income data. Observed log pay (ignoring mean-zero income shock $\xi$) is:

$$\log \text{pay}_{i,c} \equiv y^I_i + y^C_c + y^x(x^I_i, x^C_c) + y^e_{i,c} \quad (23)$$

Combining equations (12) and (23) yields:

$$\log \text{pay}_{i,c} = V_{i,c} - l^x(x^I_i, x^C_c) + \frac{1}{2} \tilde{\gamma}_c \sigma^2_c - \tilde{l}_{i,c} \quad (24)$$

We can then take the expectation of log pay conditional on career $c$ having the highest $V_{i,c}$ from the equation (16):

$$E[\log \text{pay}_{i,c} \mid V_{i,c} \geq V_{i,c} \forall c] = \overline{V(x)} - l^x(x^I_i, x^C_c) + \frac{1}{2} \tilde{\gamma}_c \sigma^2_c - E[\tilde{l}_{i,c} \mid V_{i,c} > V_{i,c} \forall c]$$

$$= \mu + \beta \gamma_{em} + y^I_i + \beta \ln\left(\sum_q s_q e(y^C_q + v(x^I_i, x^C_c) - \frac{1}{2} \tilde{\gamma}_c \sigma^2_q) / \beta \right)$$

$$- l^x(x^I_i, x^C_c) + \frac{1}{2} \tilde{\gamma}_c \sigma^2_c - E[\tilde{l}_{i,c} \mid V_{i,c} > V_{i,c} \forall c] \quad (25)$$

Next, we take the expectation of $V_{i,c}$ from equation (12) conditional on career $c$ having the highest $V_{i,c}$:

$$E[V(i, c) \mid V_{i,c} \geq V_{i,c} \forall c] = \overline{V(x)} = y^I_i + y^C_c + v(x^I_i, x^C_c) - \frac{1}{2} \tilde{\gamma}_c \sigma^2_c$$

$$+ E[y^e_{i,c} + \tilde{l}_{i,c} \mid V_{i,c} \geq V_{i,c} \forall c] \quad (26)$$
Plugging equation (16) into equation (26) and re-arranging terms yields:

\[
E[y_{i,c} + \tilde{l}_{i,c} | V_{i,c} \geq V_{i,c'} \forall c'] = \mu + \beta \gamma_{em} + \beta \ln \left( \sum_q s_q e^{(y^C_q + v(x^I_i, x^C_i) - \frac{1}{2} \tilde{\gamma}_i \sigma^2_q)/\beta} \right) - y^C_c - v(x^I_i, x^C_c) + \frac{1}{2} \tilde{\gamma}_i \sigma^2_c. \tag{27}
\]

By assuming joint normality of \(y_{i,c}^\varepsilon\) and \(\tilde{l}_{i,c}^\varepsilon\), \(E[y_{i,c} + \tilde{l}_{i,c} + \tilde{l}_{i,c}^\varepsilon | V_{i,c} \geq V_{i,c'} \forall c']\) from equation (27) identifies \(E[\tilde{l}_{i,c}^\varepsilon | V_{i,c} > V_{i,c'} \forall c']\) in equation (25):

\[
E[\tilde{l}_{i,c}^\varepsilon | V_{i,c} \geq V_{i,c'} \forall c'] = \frac{\text{var}(\tilde{l}_{i,c}^\varepsilon)}{\text{var}(y_{i,c}^\varepsilon + \tilde{l}_{i,c}^\varepsilon)} \left( \mu + \beta \gamma_{em} + \beta \ln \left( \sum_q s_q e^{(y^C_q + v(x^I_i, x^C_i) - \frac{1}{2} \tilde{\gamma}_i \sigma^2_q)/\beta} \right) - y^C_c - v(x^I_i, x^C_c) + \frac{1}{2} \tilde{\gamma}_i \sigma^2_c. \right. \tag{28}
\]

Plugging equation (28) into equation (25) yields:

\[
E[\log \text{pay}_{i,c} | V_{i,c} \geq V_{i,c'} \forall c'] = (\mu + \beta \gamma_{em})(1 - \frac{\text{var}(\tilde{l}_{i,c}^\varepsilon)}{\text{var}(y_{i,c}^\varepsilon + \tilde{l}_{i,c}^\varepsilon)}) + y^I_c + \beta \ln \left( \sum_q s_q e^{(y^C_q + v(x^I_i, x^C_i) - \frac{1}{2} \tilde{\gamma}_i \sigma^2_q)/\beta} \right)(1 - \frac{\text{var}(\tilde{l}_{i,c}^\varepsilon)}{\text{var}(y_{i,c}^\varepsilon + \tilde{l}_{i,c}^\varepsilon)}) + y^C_c \frac{\text{var}(\tilde{l}_{i,c}^\varepsilon)}{\text{var}(y_{i,c}^\varepsilon + \tilde{l}_{i,c}^\varepsilon)} - v(x^I_i, x^C_c)(1 - \frac{\text{var}(\tilde{l}_{i,c}^\varepsilon)}{\text{var}(y_{i,c}^\varepsilon + \tilde{l}_{i,c}^\varepsilon)}) + \frac{1}{2} \tilde{\gamma}_i \sigma^2_c (1 - \frac{\text{var}(\tilde{l}_{i,c}^\varepsilon)}{\text{var}(y_{i,c}^\varepsilon + \tilde{l}_{i,c}^\varepsilon)}). \tag{29}
\]

The first line in this log pay equation depends on neither individual attributes \((X^I)\) nor chosen career attributes \((X^C)\). The second line depends on individual attributes \((X^I, \text{ specifically } y^I_c \text{ and } \gamma_i)\) but not chosen career attributes. The third line depends on chosen career attributes (specifically \(y^C_c \text{ and } x^{CO}_c\) but not \(\sigma^2_c\)) and individual observables \((x^{IO}_i)\). The final line depends on both an individual attribute \((\gamma_i)\) and a career attribute \((\sigma^2)\).

Equation (29) suggests a simple regression that can be used to recover the rel-
ative importance of taste shocks ($l_{i,c}^{\varepsilon}$) compared with all shocks ($y_{i,c}^{\varepsilon} + l_{i,c}^{\varepsilon}$). The regression predicts pay with (numbers match the line number in equation (29)): 1) a constant; 2) individual-specific controls (including risk aversion); 3) career-attribute controls (particularly a measure of income risk); and, 4) the interaction of individual- and career-specific controls (besides risk aversion and income risk). This fourth line in equation (29) shows that the coefficient interaction between risk aversion ($\gamma_i$) and income risk ($\gamma_i$) identifies $\frac{1}{2} \times \left(1 - \frac{\text{var}(l_{i,c}^{\varepsilon})}{\text{var}(y_{i,c}^{\varepsilon} + l_{i,c}^{\varepsilon})}\right)$.

If people dislike risk, they must be compensated in some way for taking on more of it. The more risk-averse a person is, the greater such compensation must be. This compensation could come in the form of higher pay or more career enjoyment. Risk-averse people will only choose risky jobs if they love them or are very productive in them (thereby earning particularly high pay). In a world in which most idiosyncratic variation is in enjoyment, we will see risk-averse people compensated by choosing risky jobs they particularly enjoy. In a world in which most idiosyncratic variation is in ability or productivity, we will see risk-averse people compensated by choosing jobs at which they particularly excel and therefore earn higher pay. We should not see this pattern among the risk-neutral.

3 Data

Our data are the core sample of the Panel Study of Income Dynamics (PSID). The PSID was designed as a nationally representative panel of U.S. households (Hill, 1991); it provides annual or biennial labor income spanning the years 1968 to 2005. Restricting ourselves male household heads aged 22 through 60 gives us 52,181 observations on 3,041 individuals with 17 years of recorded data per individual.

Table 1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>st. dev.</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>age (years)</td>
<td>42.9</td>
<td>7.8</td>
<td>28</td>
<td>60</td>
</tr>
<tr>
<td>education (years)</td>
<td>14.1</td>
<td>2.2</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>married</td>
<td>84.8%</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>black</td>
<td>3.8%</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>annual income (2005 $s)</td>
<td>$57,756</td>
<td>$52,990</td>
<td>0</td>
<td>$753,233</td>
</tr>
<tr>
<td>family size</td>
<td>3.2</td>
<td>1.3</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td># of obs.</td>
<td>1,427</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

Table 1 summarizes 1997 data for the 1,427 male household heads used in estimating coefficient values in equation (22). We restrict the sample to the 1,427 male household heads with income data in all years 1991-1997 at ages 22 to 60, with non-missing volatility and risk tolerance values in 1997, and non-missing Jensen-Shore volatility values over 1991-1996.

on average. As mentioned previously, we focus on excess log income as our outcome measure, which is the residual from a regression to predict the natural log of labor income. Furthermore, we restrict the sample to individuals with income (and therefore volatility) values in 1991 through 1996, with risk tolerance responses recorded in the 1997 wave. There are 1,429 individuals like this. Summary statistics about the demographics of this group in 1997 are shown in Table 1.

3.1 Income Volatility

Using data from the PSID, we calculate two “off-the-shelf” measures of income volatility. Jensen and Shore develop a methodology to estimate non-parametrically the distribution of volatility of excess log income as our outcome measure, which is the residual from a regression to predict the natural log of labor income. This regression is weighted by PSID-provided sample weights, normalized so that the average weight in each year is the same. We use the following as covariates in this regression: a cubic in age for each level of educational attainment (none, elemen-
tary, junior high, some high school, high school, some college, college, graduate school); the presence and number of infants, young children, and older children in the household; the total number of family members in the household, and dummy variables for each calendar year. Including calendar year dummy variables eliminates the need to convert nominal income to real income explicitly.\(^{17}\)

While some other papers have dropped observations with missing and zero income (Gottschalk and Moffitt, 2002) or modeled unemployment explicitly (Pistaferri, 2002), neither route is available to us because the method in Jensen and Shore is not designed to handle missing data or zeros. Instead, Jensen and Shore fill in hot-deck imputed missing values when calculating volatility. Aside from using their volatility values, we do not explicitly use bootstrapped income data. We follow Jensen and Shore in using top- and bottom-codes. The income dynamics used by Jensen and Shore to estimate income volatility are quite standard, characterizing the evolution of excess log income for individual \(i\) over time \(t\) (Carroll and Samwick, 1997; Meghir and Pistaferri, 2004). Excess log income \(y_{i,t}\) is modeled as the sum of permanent income, transitory income, and error \(e_{i,t}\).

\[
y_{i,t} = \sum_{k=1}^{t-3} \omega_{i,k} + \sum_{k=t-2}^{t} \phi_{\omega,t-k} \cdot \omega_{i,k} + \sum_{k=t-2}^{t} \phi_{\varepsilon,t-k} \cdot \varepsilon_{i,k} + e_{i,t}
\]

Permanent income is the weighted sum of past permanent shocks \(\omega_{i,k}\) to income. Transitory income is the weighted sum of recent transitory shocks \(\varepsilon_{i,k}\) to income.\(^{18}\)

The permanent shock, transitory shock, and error term are assumed to be nor-

17 Working with excess log real income is also standard in this literature (Meghir and Pistaferri, 2004; Carroll and Samwick, 1997).

18 In this framework model, permanent shocks come into effect over three periods and transitory shocks fade completely after three periods, giving us three permanent weight parameters \((\phi_{\omega,0}, \phi_{\omega,1}, \phi_{\omega,2})\) and three transitory weight parameters \((\phi_{\varepsilon,0}, \phi_{\varepsilon,1}, \phi_{\varepsilon,2})\). We refer to these weights \(\phi\) collectively as the income process parameters, which will need to be estimated in our model. Jensen and Shore posit flat prior distributions for each weight parameter (i.e. \(p(\phi) \propto 1\)). However, in order to give meaning to the magnitude of our transitory shocks, we normalize the weights placed on transitory shocks to sum to one \((\sum_k \phi_{\varepsilon,k} = 1)\)
mally distributed as well as independent of one another over time and across individuals. The permanent shocks \( \omega_{i,t} \) have mean zero and variance \( \sigma_{\omega,i,t}^2 \equiv \mathbb{E}[\omega_{i,t}^2] \); the transitory shocks \( \varepsilon_{i,t} \) have mean zero and variance \( \sigma_{\varepsilon,i,t}^2 \equiv \mathbb{E}[\varepsilon_{i,t}^2] \):

\[
\begin{pmatrix}
\omega_{i,t} \\
\varepsilon_{i,t}
\end{pmatrix}
\sim
\mathcal{N}
\begin{pmatrix}
0 \\
0
\end{pmatrix},
\begin{bmatrix}
\sigma_{\omega,i,t}^2 & 0 \\
0 & \sigma_{\varepsilon,i,t}^2
\end{bmatrix}
\]

(31)

We refer to \( \sigma_{i,t}^2 \equiv (\sigma_{\varepsilon,i,t}^2, \sigma_{\omega,i,t}^2) \) jointly as the volatility parameters. Finally, we have “noise variance” which refers to the variance of measurement error \( \gamma^2 \equiv \mathbb{E}[\varepsilon_{i,t}^2] \) that is constant across individuals and over time.

Jensen and Shore (2009a,b) develop a Markovian hierarchical Dirichlet Process (MHDP) prior that they use to estimate the distribution of ex-ante expected volatility. We use the estimates of the ex-ante expected permanent volatility distribution from their paper.

As an alternative measure, we use the estimator for the variance of permanent income changes proposed by Meghir and Pistaferri (2004):

\[
\sigma_{\omega,i,t}^2 = \mathbb{E}[(y_{i,t} - y_{i,t-1}) \times (y_{i,t+m} - y_{i,t-1-n})]
\]

(32)

This paper uses 1991-1996 averages of estimates of permanent volatility obtained using the Jensen and Shore (2009a,b) methodology, and also using the Meghir and Pistaferri (2004) moment. Note that the latter is much more dispersed, as it measures the distribution of ex-post volatility, not ex-ante expected volatility as Jensen and Shore (2009a,b) aim to do.

### 3.2 Risk-Aversion

In 1996, the PSID included a series of survey questions which aimed to elicit estimates of risk tolerance, \( 1/\gamma \). Respondents were asked whether they would be
willing to take a series of lotteries which varied in compensation for risk. Based on which gambles the respondents were and were not willing to take, risk tolerance was identified as within 4 ranges.

Table 2 presents the joint distribution of volatility and risk aversion. In that table, \( \sigma^2 \) values are divided into 10 bins. These are formed using \( \sigma \) intervals of 0.01, grouping such intervals together until each grouping contained at least 0.5\% of the \( \sigma \) values. The distribution of \( \sigma^2 \) values is also shown in the upper-left panel of Figure 2; this distribution is broken into the coarser bins in the lower-left panel. The upper-right panel shows the proportion of the data in each of these 10 bins; the upper-right and lower-left panels are identical except that their axes are scaled differently. Note that individuals with high risk aversion were slightly less likely to have the highest volatility values. This is also apparent in reduced-form regressions to predict volatility with risk-aversion and other controls. These are shown in Table 3. In both Table 2 and Table 3, the top panel uses the Jensen and Shore volatility moment while the bottom panel uses the Meghir and Pistaferri moment.

Table 4 shows (1991-1996) income volatility and (1997) risk-aversion data by (1991) “one-digit” occupational category. Note that, while income volatility varies across occupations, the correlation between occupational income volatility and occupational risk tolerance is quite low.

4 Estimation

If we could observe the joint distribution of data \( \{\sigma^2, \tilde{\gamma}, x^{IO}, x^{CO}\} \), then estimation of equations (21) and (22) by maximum likelihood is straightforward. We need only choose a parametric (or nonparametric) structure for \( f(\sigma^2 | \tilde{\gamma} = 0, x^{IO}, x^{CO}) \), and estimate its parameters along with \( \beta \) by maximum likelihood. Table 2 shows
Table 2: Estimated distribution of income volatility by self-reported risk-aversion

<table>
<thead>
<tr>
<th>$\sigma^2 &gt; \tilde{\sigma}^2 \leq \tilde{\sigma}^2$</th>
<th>unconditional $\hat{\sigma}^2$ distribution</th>
<th>$\sigma^2$ distribution conditional on $\tilde{\gamma}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[0.2)</td>
<td>[2, 3.84)</td>
</tr>
<tr>
<td>0.14$^2$ 0.15$^2$</td>
<td>27.55 %</td>
<td>24.95 %</td>
</tr>
<tr>
<td>0.15$^2$ 0.16$^2$</td>
<td>54.78 %</td>
<td>49.50 %</td>
</tr>
<tr>
<td>0.16$^2$ 0.17$^2$</td>
<td>6.04 %</td>
<td>6.61 %</td>
</tr>
<tr>
<td>0.17$^2$ 0.18$^2$</td>
<td>2.15 %</td>
<td>3.46 %</td>
</tr>
<tr>
<td>0.18$^2$ 0.20$^2$</td>
<td>2.31 %</td>
<td>4.52 %</td>
</tr>
<tr>
<td>0.20$^2$ 0.40$^2$</td>
<td>3.41 %</td>
<td>4.56 %</td>
</tr>
<tr>
<td>0.40$^2$ 0.60$^2$</td>
<td>1.22 %</td>
<td>1.97 %</td>
</tr>
<tr>
<td>0.60$^2$ 0.80$^2$</td>
<td>1.61 %</td>
<td>2.65 %</td>
</tr>
<tr>
<td>0.80$^2$ 1.00$^2$</td>
<td>0.22 %</td>
<td>0.00 %</td>
</tr>
<tr>
<td>1.00$^2$ 1.00$^2$</td>
<td>0.71 %</td>
<td>1.78 %</td>
</tr>
</tbody>
</table>

Mean $\tilde{\sigma}^2$ | 0.05 | 0.07 | 0.04 | 0.05 | 0.04 |
St. Dev. $\tilde{\sigma}^2$ | 0.15 | 0.23 | 0.11 | 0.16 | 0.11 |
# of Obs. | 1,427 | 303 | 229 | 262 | 633 |
% of Obs. | 100 % | 21.23 % | 16.05 % | 18.36 % | 44.36 % |

<table>
<thead>
<tr>
<th>$\sigma^2 &gt; \tilde{\sigma}^2$</th>
<th>unconditional $\hat{\sigma}^2$ distribution</th>
<th>$\sigma^2$ distribution conditional on $\tilde{\gamma}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[0.2)</td>
<td>[2, 3.84)</td>
</tr>
<tr>
<td>-1.00 -0.50</td>
<td>2.46 %</td>
<td>2.11 %</td>
</tr>
<tr>
<td>-0.50 -0.25</td>
<td>2.43 %</td>
<td>2.09 %</td>
</tr>
<tr>
<td>-0.25 -0.05</td>
<td>6.56 %</td>
<td>7.99 %</td>
</tr>
<tr>
<td>-0.05 0.00</td>
<td>21.34 %</td>
<td>19.67 %</td>
</tr>
<tr>
<td>0.00 0.03</td>
<td>31.88 %</td>
<td>24.75 %</td>
</tr>
<tr>
<td>0.03 0.05</td>
<td>10.15 %</td>
<td>8.31 %</td>
</tr>
<tr>
<td>0.05 0.10</td>
<td>8.01 %</td>
<td>9.50 %</td>
</tr>
<tr>
<td>0.10 0.40</td>
<td>10.54 %</td>
<td>11.33 %</td>
</tr>
<tr>
<td>0.40 1.00</td>
<td>4.95 %</td>
<td>10.96 %</td>
</tr>
<tr>
<td>1.00 1.00</td>
<td>1.67 %</td>
<td>3.29 %</td>
</tr>
</tbody>
</table>

Mean $\tilde{\sigma}^2$ | 0.04 | 0.08 | 0.04 | 0.03 | 0.04 |
St. Dev. $\tilde{\sigma}^2$ | 0.35 | 0.47 | 0.22 | 0.28 | 0.31 |
# of Obs. | 1,426 | 303 | 228 | 262 | 633 |
% of Obs. | 100 % | 21.25 % | 15.99 % | 18.37 % | 44.36 % |

Table 2 shows the distribution of $\sigma^2$ estimates, unconditionally and by self-reported risk-aversion bin. The estimates of $\sigma^2$ are the average of 1991-1996 permanent volatility estimates using the Jensen and Shore (2010) methodology and the average of 1992-1995 permanent volatility estimates using the Meghir and Pistaferri (2004) moment. Jensen-Shore values are top-coded at 1; Meghir-Pistaferri values are top and bottom coded at 1 and -1, respectively. Self-reported risk-aversion ranges are from responses to the 1996 risk-tolerance supplement to the PSID.
Figure 2: Jensen-Shore Permanent Income Volatility Bins

Distribution of Income Volatility
(many equally spaced bins)

Distribution of Income Volatility
(fraction of data in each of 10 bins)

The upper-left panel presents the distribution of 1991 – 1996 $\sigma$ estimates from Jensen and Shore, a histogram of the standard deviation of permanent income changes. The lower-left panel shows how we collect these values into 10 bins. The upper-right panel is identical to the lower-left in showing the fraction of individuals in each $\sigma$ bin; it differs only in the x-axis, which shows bin instead of $\sigma$ values. The lower-right panel presents estimates of $f(\sigma|\tilde{\gamma} = 0)$. These are normalized by dividing by the value in the upper-right panel and subtracting one. This shows the degree to which risk-neutral individuals are estimated to over-weight or under-weight this bin relative the population as a whole. This panel shows 95% confidence intervals from a likelihood ratio test (where only this probability but no other parameters are restricted).
Table 3: Relationship between Income Volatility (’91-’96) and Risk Aversion (’97)

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>Jenson-Shore</th>
<th>Meghir-Pistaferri</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[\tilde{\gamma} \mid \text{bin}]$</td>
<td>Income Risk Level</td>
<td>Income Risk Log</td>
</tr>
<tr>
<td>-0.006</td>
<td>-0.007</td>
<td>-0.015</td>
</tr>
<tr>
<td>(0.003)*</td>
<td>(0.003)*</td>
<td>(0.012)</td>
</tr>
<tr>
<td>age</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.003)*</td>
</tr>
<tr>
<td>controls</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.006</td>
<td>0.009</td>
</tr>
<tr>
<td># of Obs.</td>
<td>1,427</td>
<td>1,427</td>
</tr>
</tbody>
</table>

Table 3 shows the OLS regressions to predict individual-specific measures of income risk with self-reported risk-aversion bin. $E[\tilde{\gamma} \mid \text{bin}]$ refers to the expected value of risk aversion conditional on risk aversion bin, which we estimate using the signal-noise structure identified in Kimball, Sahm, and Shapiro. * Indicates significance at the 5% level, † indicates significance at the 1% level.

We observe true log risk tolerance ($\ln(1/\tilde{\gamma})$) plus noise ($e$), placed into bins, so that...
Table 4: Distribution of income risk and risk-aversion by broad occupational category

<table>
<thead>
<tr>
<th>occupation</th>
<th>obs</th>
<th>Jenson-Shore $\sigma^2$ mean</th>
<th>st.dev</th>
<th>Meghir-Pistaferri $\sigma^2$ mean</th>
<th>st.dev.</th>
<th>$\gamma$ bin least ← risk averse → most</th>
</tr>
</thead>
<tbody>
<tr>
<td>prof./tech.</td>
<td>329</td>
<td>0.031</td>
<td>0.062</td>
<td>0.039</td>
<td>0.174</td>
<td>20 %</td>
</tr>
<tr>
<td>managers</td>
<td>320</td>
<td>0.073</td>
<td>0.237</td>
<td>0.064</td>
<td>0.476</td>
<td>27 %</td>
</tr>
<tr>
<td>clerical</td>
<td>210</td>
<td>0.035</td>
<td>0.116</td>
<td>0.049</td>
<td>0.332</td>
<td>20 %</td>
</tr>
<tr>
<td>craftsmen</td>
<td>269</td>
<td>0.049</td>
<td>0.120</td>
<td>0.028</td>
<td>0.294</td>
<td>16 %</td>
</tr>
<tr>
<td>operators</td>
<td>151</td>
<td>0.039</td>
<td>0.077</td>
<td>0.023</td>
<td>0.245</td>
<td>19 %</td>
</tr>
<tr>
<td>laborers</td>
<td>50</td>
<td>0.033</td>
<td>0.055</td>
<td>0.050</td>
<td>0.161</td>
<td>14 %</td>
</tr>
<tr>
<td>farmers</td>
<td>34</td>
<td>0.163</td>
<td>0.352</td>
<td>0.008</td>
<td>0.530</td>
<td>18 %</td>
</tr>
<tr>
<td>n/a</td>
<td>52</td>
<td>0.058</td>
<td>0.111</td>
<td>0.166</td>
<td>0.397</td>
<td>33 %</td>
</tr>
<tr>
<td>overall</td>
<td>1,427</td>
<td>0.050</td>
<td>0.152</td>
<td>0.048</td>
<td>0.334</td>
<td>22 %</td>
</tr>
</tbody>
</table>

Table 4 shows the distribution of self-reported risk preference by one-digit occupational categories. The n/a category includes non-responses.

A given observation lies in a given bin if $\ln(1/\bar{\gamma}) > \text{bin}$ and $\ln(1/\bar{\gamma}) < \text{bin}$. Again, Table 2 shows these ranges and the fraction of observed data that falls into each.\(^{19}\)

We can then identify the relationship between our data ($f(\sigma^2 | 1/\ln(\bar{\gamma}) \text{ bin})$) and the object we wish to estimate ($f(\sigma^2 | \bar{\gamma})$) from equation (22):

\[
 f(\sigma^2 | 1/\ln(\bar{\gamma}) \text{ bin}) = \int_{\bar{\gamma}} f(\sigma^2 | \bar{\gamma}) f_{\ln(1/\bar{\gamma})}(\ln (1/\bar{\gamma}) | \text{ bin}) \, d\bar{\gamma} 
\]

\[ (35) \]

\[
 f(\ln (1/\bar{\gamma}) | 1/\ln(\bar{\gamma}) \text{ bin}) = f_{\ln(1/\bar{\gamma})}(\ln (1/\bar{\gamma})) \cdot \frac{\text{pr}(1/\ln(\bar{\gamma}) \text{ bin} | \bar{\gamma})}{\text{pr}(1/\ln(\bar{\gamma}) \text{ bin})} 
\]

Given the distribution of true variation and classical measurement error estimated by Kimball, Sahm, and Shapiro, it is trivial to calculate $f_{\ln(1/\bar{\gamma})}(\ln (1/\bar{\gamma}))$ and $\text{pr}(\text{ bin} | \bar{\gamma})$ for each $\bar{\gamma}$ in our grid for each of the four risk-aversion bins; $\text{pr}(\text{ bin})$ is similarly easy to calculate for each of the four risk-aversion bins.

Armed with this distribution of $\bar{\gamma}$, we search for maximum likelihood estimates

\(^{19}\)We approximate this distribution with a 38 element grid, assigning a probability that $\bar{\gamma}$ will be each of the following values: \{0.5, 1.25, 2, 2.5, 3, 3.4, 3.8, 4.5, 5.5, ..., 9.5, 10, 10.5, 11, 12, ..., 34\}.
of \( f(\sigma^2 \mid \tilde{\gamma} = 0) \) and \( \beta \) iteratively. First, we guess values of \( f(\sigma^2 \mid \tilde{\gamma} = 0) \) and \( \beta \). Next, we calculate \( f(\sigma^2 \mid \tilde{\gamma}) \) for each value of \( \sigma^2 \) and \( \tilde{\gamma} \) on our grid. Next, we calculate \( f(\sigma^2 \mid \tilde{\gamma}) \) for each of the 10 grid values of \( \sigma^2 \) and each of the four coarse bins for \( \tilde{\gamma} \) by integrating over each value of \( \tilde{\gamma} \) possible in each bin. This gives the likelihood of an observation lying in one of the \( 10 \times 4 = 40 \) possible ranges we observe in Table 2. We then compute the likelihood of observing the data in Table 2. We search over \( f(\sigma^2 \mid \tilde{\gamma} = 0) \) and \( \beta \) to find values that maximizes the likelihood.

5 Results

Equations (21) and (22) present the key model parameters we estimate: \( \beta \) and \( f(\sigma^2 \mid \tilde{\gamma} = 0, x^{IO}, x^{CO}) \). Recall that \( \beta \) is proportional to the standard deviation of idiosyncratic individual-career-specific taste and skill values; \( f(\sigma^2 \mid \tilde{\gamma} = 0, x^{IO}, x^{CO}) \) is the distribution of income risk for the careers chosen by a risk-neutral worker.

These results are shown in Table 5. The table includes two panels, the left one showing results using Jensen and Shore measures of average permanent volatility; the right one showing results using the Meghir and Pistaferri permanent volatility moment. In each panel, the first column presents results without additional covariates; the second presents results with years of education (linear), age (linear) and race (dummy) controls; the third presents results which also add one-digit occupation category controls.

The \( \beta \) point estimate using the Jensen and Shore volatility without other controls is 0.76, so that the standard deviation of idiosyncratic career values is \( 96\% \times \pi/\sqrt{6} \) of income (in log points). Using a likelihood ratio test, the 95% confidence interval for \( \beta \) is \( 0.36 < \beta < 2.86 \). Although these estimates are large, we view the lower-bound on \( \beta \) as entirely plausible; it implies a dispersion of idiosyncratic taste or skill of 46% of income. Said differently, increasing by one standard deviation the
idiosyncratic career skill or taste is equivalent to a pay increase of 46%. Adding controls for age, race, education and occupation changes the point estimate and upper/lower bounds of $\beta$ only slightly.

Table 5 also shows estimates generated from (Meghir and Pistaferri, 2004) volatility moments. With a point estimate of 3.14 without covariates (compared with 0.76 for Jensen and Shore, and broadly similar with additional covariates), these estimates are substantially larger. This is consistent with the attenuation bias we would expect in estimates of $1/\beta$ using Meghir and Pistaferri. The volatility values from Meghir and Pistaferri are more dispersed because they capture the distribution of volatility estimates (which include error in the measurement of volatility); by contrast, Jensen and Shore aim to estimate the distribution of expected volatility.

This can be seen graphically in the lower panels of Figure 2; the lower-left one shows results for Jensen-Shore income volatility measures; the lower-right one shows results for Meghir-Pistaferri ones. These figures show $f(\sigma^2 | \tilde{\gamma} = 0)/f(\sigma^2) - 1$, the degree to which risk-neutral workers over-weight or under-weight careers within a given income risk range, when the model is estimated without covariates. Figures also show 95% confidence intervals around these estimates (separately for each $\sigma^2$ bin) from a likelihood ratio test. The upward slope in point estimates indicates that risk-neutral people modestly under-weight (unconditionally common) low-risk careers but substantially over-weight (unconditionally rare) high-risk ones.

Equation (22) shows that the distribution of $\sigma^2$ choices by risk-neutral people may reflect the distribution of career options $f^C$ or the relative value of those options ($E \left[ (y^C + v (x^I, x^C))^\frac{1}{\beta} | \sigma^2 \right]$). There is no way to differentiate these options without a model of wage adjustment. At one extreme, we can assume that employers’ supply of careers is completely elastic, so that there is no compensation.
Table 5: Parameter Estimates

<table>
<thead>
<tr>
<th></th>
<th>Jensen-Shore σ</th>
<th>Meghir-Pistaferri σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$</td>
<td>0.756 0.659 0.770</td>
<td>3.137 2.956 3.815</td>
</tr>
<tr>
<td>$\beta &lt;$</td>
<td>0.359 0.328 0.360</td>
<td>?? ?? ??</td>
</tr>
<tr>
<td>$\beta &gt;$</td>
<td>2.854 2.194 3.199</td>
<td>?? ?? ??</td>
</tr>
<tr>
<td>lowest income</td>
<td>24.4 % 28.0 % 28.9 %</td>
<td>1.0 % 3.0 % 2.1 %</td>
</tr>
<tr>
<td>highest income</td>
<td>50.1 % 56.0 % 56.9 %</td>
<td>1.7 % 3.0 % 2.4 %</td>
</tr>
<tr>
<td>risk ↑</td>
<td>5.7 % 6.2 % 6.2 % 2.2 %</td>
<td>20.1 % 22.4 % 22.1 %</td>
</tr>
<tr>
<td>$f(\sigma</td>
<td>\gamma = 0)$</td>
<td>4.6 % 3.3 % 2.7 %</td>
</tr>
<tr>
<td>↑</td>
<td>2.4 % 2.4 % 2.3 %</td>
<td>31.0 % 32.8 % 32.0 %</td>
</tr>
<tr>
<td>↓</td>
<td>2.5 % 0.9 % 0.5 %</td>
<td>8.2 % 8.0 % 8.1 %</td>
</tr>
<tr>
<td>highest income</td>
<td>4.5 % 0.8 % 0.3 %</td>
<td>12.1 % 9.3 % 10.5 %</td>
</tr>
<tr>
<td>highest risk</td>
<td>0.8 % 0.1 % 0.0 %</td>
<td>7.4 % 3.1 % 4.3 %</td>
</tr>
<tr>
<td>age × $\sigma^2$</td>
<td>. 0.085 0.08</td>
<td>. -0.005 -0.008</td>
</tr>
<tr>
<td>edu. × $\sigma^2$</td>
<td>. -0.006 0.21</td>
<td>. 0.111 0.108</td>
</tr>
<tr>
<td>race × $\sigma^2$</td>
<td>no yes yes</td>
<td>no yes yes</td>
</tr>
<tr>
<td>occ. × $\sigma^2$</td>
<td>no no yes</td>
<td>no no yes</td>
</tr>
</tbody>
</table>

Table 5 displays the estimates of $\beta$ (including a 95% confidence interval), $f(\sigma^2|\gamma = 0)$, and $\theta$ from equations (21) and (22). The point estimates of $\beta$ correspond to the variance of idiosyncratic taste and skill shocks. $f(\sigma^2|\gamma = 0)$ is the probability that a risk-neutral individual populates each of the 10 $\sigma^2$ bins. The vector $\theta$ represents the coefficient estimates of these controls.

for income risk; in this case, risky career options are merely chosen less frequently than safe ones. At the other extreme, we can assume that employers’ supply of careers is completely inelastic, so that wages adjust until the unconditional distribution of chosen careers ($f(\sigma^2)$) is equal to the distribution of career options ($f^C$). Wages adjust until all career paths are filled.

In the case of inelastic career supply, we implicitly observe $f^C$ and can identify $y^C + v(x^I, x^C)$, the premia profile needed to fill all careers. Assuming no heterogeneity conditional on $\sigma^2$, equation (22) implies that $y^C + v(x^I, x^C) = (\frac{f(\sigma^2|\gamma = 0)}{f^C})^\beta$. Normalized estimates of this risk-return menu are shown in the solid line in Figure.

31
Figure 3 shows the estimated income premium $y^C$ at the midpoint of each $\sigma^2$ bin. The dashed curve reflects the perfect-sorting case from equation (14). The solid-curve reflects the required risk-premium needed to rationalize the data, under the assumption that career supply is inelastic so that pay adjusts so that the income risk distribution of career options equals the income risk distribution of chosen careers, when equation (21) is estimated without covariates.

3. Note the substantial risk premium required for high income risk bins. This is consistent with the idea that the marginal person choosing a risky career in our model is not very risk tolerant; while she is typically only slightly more risk tolerant than average, she chooses a high-risk career because she happens to love it or excel at it. In equilibrium, these risk-averse people must be offered significant compensation to undertake risky careers.

We can contrast this with the hypothetical risk-return profile if $\beta = 0$. Absent idiosyncratic taste and skill, we would see perfect sorting of the most risk tolerant people into the riskiest careers. As a result, risk premia for high-risk careers could be much lower.

Next, we use income data to decompose $\varepsilon_{i,c}$ into idiosyncratic career skill ($y^C_{i,c}$) and taste ($\widetilde{l}_{i,c}$). We estimate equation (29) by OLS to identify $(1 - \text{var}(\widetilde{l}_{i,c})/\text{var}(y^C_{i,c} + \widetilde{l}_{i,c}))$, the coefficient on $\frac{1}{2} \times \sigma^2 \times \gamma$. If this coefficient is $< .5$, the variation in career choice is mostly in idiosyncratic taste rather than skill. The intuition here is that
risk-averse people demand a larger “compensation” to enter high-risk careers; that compensation may come in the form of higher productivity (and therefore higher pay) or in higher career enjoyment. As a result, the gap in compensation between high- and low-risk careers will be greatest for those with the highest risk-aversion. If we do not observe a pay gap, this compensation must be in the form of idiosyncratic taste (loving your job). Table 6 shows the results from regressions to predict pay with $\gamma$, $\sigma^2$, their interaction, and covariates. While the coefficient on $\frac{1}{2} \times \sigma^2 \times \gamma$ is noticeably smaller than 0.5, it is not statistically significant. However, we can reject the hypothesis that the variation in career choice is due mostly to skill rather than taste. Thus, we can conclude that people vary widely in how much they think they would like various jobs, not in how good they would be at those jobs. Finally, note the similarity between this regression and the risk-augmented Mincer equations from Hartog (2009), which provides a consistency check on our particular sample.

\footnote{Note that $\gamma$ refers to $E[\gamma | \tilde{\gamma}_{\text{bin}}]$, which is based on the distribution of $\tilde{\gamma}$ and measurement error proposed by Kimball, Sahm, and Shapiro (2009).}
Table 6: Impact of income risk and risk-aversion on income

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Log Average Income: Jenson-Shore</th>
<th>Log Average Income: Meghir-Pistaferri</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>-0.176 (0.139)</td>
<td>-0.077 (0.065)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>-0.257 (0.123)*</td>
<td>-0.098 (0.058)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>-0.050 (0.310)</td>
<td>-0.132 (0.137)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>-0.132 (0.275)</td>
<td>-0.096 (0.122)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>0.959 (0.868)</td>
<td>-0.338 (0.152)*</td>
</tr>
<tr>
<td>$\ln(\sigma^2)$</td>
<td>-0.050 (0.137)</td>
<td>0.098 (0.065)</td>
</tr>
<tr>
<td>$\gamma = 2^{nd}$</td>
<td>0.031 (0.073)</td>
<td>0.031 (0.073)</td>
</tr>
<tr>
<td>lowest</td>
<td>0.014 (0.064)</td>
<td>0.017 (0.064)</td>
</tr>
<tr>
<td>$\gamma = 2^{nd}$</td>
<td>0.118 (0.064)</td>
<td>0.109 (0.064)</td>
</tr>
<tr>
<td>highest</td>
<td>0.089 (0.063)</td>
<td>0.087 (0.063)</td>
</tr>
<tr>
<td>$\gamma = 2^{nd}$</td>
<td>0.070 (0.072)</td>
<td>0.071 (0.072)</td>
</tr>
<tr>
<td>highest</td>
<td>0.118 (0.064)</td>
<td>0.109 (0.071)</td>
</tr>
<tr>
<td>$\sigma^2 \times \gamma$</td>
<td>0.083 (0.064)</td>
<td>0.071 (0.064)</td>
</tr>
<tr>
<td>$\ln(\sigma^2)$</td>
<td>0.174 (0.063)</td>
<td>0.087 (0.063)</td>
</tr>
<tr>
<td>$\ln(\sigma^2)$</td>
<td>0.085 (0.064)</td>
<td>0.087 (0.064)</td>
</tr>
<tr>
<td>age</td>
<td>no yes no yes yes yes</td>
<td>no yes no yes yes yes yes</td>
</tr>
<tr>
<td>family size</td>
<td>no yes no yes yes yes</td>
<td>no yes no yes yes yes yes</td>
</tr>
<tr>
<td>education</td>
<td>no yes yes yes yes yes yes</td>
<td>no yes yes yes yes yes yes</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.00 0.23 0.01 0.23 0.26</td>
<td>0.00 0.23 0.01 0.23 0.24</td>
</tr>
<tr>
<td># of Obs.</td>
<td>1,421 1,421 1,421 1,421 1,421</td>
<td>1,420 1,420 1,420 1,420 1,420</td>
</tr>
</tbody>
</table>

Standard errors in parenthesis: *significant at 5% level; **significant at 1% level. All results are for OLS regressions weighted by PSID-provided sample weights. “age” indicates whether a linear age control was included; “family size” indicates whether linear controls for total family size, presence and number of babies, young children, and older children were included; “race” indicates whether “white”, “black” and “other race” controls were included; “education” indicates whether a linear years of schooling variable was included. While the full sample includes 1,429 observations, 6 of these have an income of zero throughout, and consequently a missing log income. $\sigma^2$ refers to the average of Jensen and Shore’s estimates of permanent income volatility from 1991 to 1996. The dependent variable is the log of average income, averaged over the period 1991 to 1996.
6 Conclusion

This paper has documented that those who self-identify as risk-averse are more likely to have volatile incomes. Our model of optimal career choice with idiosyncratic taste and skill gives this feature of the data an economic interpretation: perceived idiosyncratic taste and/or skill varies dramatically from one career to another. The lack of an income gap between high- and low-risk careers for risk averse people (relative to risk tolerant ones) indicates that this idiosyncratic factor is taste for one career over another, and not skill in one career over another.

Another interpretation of the data is that individuals do not “choose” careers; careers are chosen for them. This reading could have a search interpretation, in which individuals are unaware of the best jobs for them. It could have a non-market interpretation, so that individuals are assigned careers or can only choose from a very narrow set; some choose from a safe set of career options while others choose from a risky set. Finally, it could have a non-optimizing interpretation, in which individuals don’t choose to maximize expected utility. In all of these readings, our results are still informative. We have identified the dollar-equivalent value of deviations from the risk-return relationship. In this way, our results could be taken as a measure of the dollar value of eliminating search frictions, the dollar value of letting people choose careers freely or from the same risk distribution, and the dollar value of more careful choice, respectively.
References


