Departamento de Economía
Facultad de Ciencias Sociales
Universidad de la República

Documentos de Trabajo

Comparative Advantage and Risk Premia in Labor Markets

German Cubas y Pedro Silos

Documento No. 02/13
Marzo 2013

ISSN 0797-7484

Comparative Advantage and Risk Premia in Labor Markets

German Cubas and Pedro Silos*

*Affiliation: Central Bank of Uruguay and FCS-University of Republic, Federal Reserve Bank of Atlanta, respectively. We would like to thank Gustavo Canavire for his excellent research assistance and, Dante Amengual, Yongsung Chang, Mark Bils, Fernando Borraz, CEPR, Christopher Carroll, Juan Dubra, Juan Carlos Hatchondo, Georgui Kambourov, David Lagakos, Martin Lopez-Daneri, Jorge Ponce, B. Ravikumar, Victor Rios-Rull, Cesare Robotti, Richard Rogerson, Yongseok Shin, Gustavo Ventura, and Ron Warren for their comments and suggestions and seminar participants at dECON FCE-UDELAR, Central Bank of Uruguay, Atlanta FED, NBER Summer Institute EFACR, St. Louis FED, the University of Iowa, and the University of Georgia. We thank the Fondo Clemente Estable of ANII (proyecto FCE-2-2011-1-6904) for their financial support. The views expressed here are those of the authors and cannot be attributed to the Central Bank of Uruguay, the Federal Reserve Bank of Atlanta or the Federal Reserve System.

Abstract

Using the Survey of Income and Program Participation (SIPP) we estimate quarterly labor earnings risk across 21 industries of the US economy. We document a significant and positive association between earnings risk (both permanent and transitory) and average log-earnings across industries. The Finance sector is 50% riskier than Government which implies a mean earnings premium of 20%. We develop an equilibrium framework to analyze the interplay between volatility in labor earnings and comparative advantage in determining the level of earnings across industries. We use the model to decompose how much of the empirical correlation represents compensation for risk and how much represents selection. The positive association between permanent risk and earnings is compensation for risk, but selection is responsible for the observed relationship between temporary risk and mean earnings.

Key words: Risk Premium, Labor Markets, Industry, Comparative Advantage
JEL Classifications: D91 · J31 · J61.

Resumen

Usando datos trimestrales de la encuesta Survey of Income and Program Participation (SIPP) estimamos el riesgo asociado a los ingresos laborales para 21 industrias de la economía de Estados Unidos. Documentamos una fuerte y positiva asociación entre el riesgo (tanto transitorio como permanente) y la media del logaritmo de los ingresos laborales entre industrias. El sector Finanzas es 50% mas riesgoso que el sector Gobierno y eso implica un premio de ingresos laborales de alrededor del 20%. Desarrollamos un modelo de equilibrio general para analizar la relación entre la volatilidad de los ingresos laborales y las ventajas comparativas de los trabajadores en la determinación del nivel de ingreso para distintas industrias. Usamos el modelo para descomponer la correlación entre compensación por riesgo y selección. La relación positiva entre riesgo permanente e ingresos es efectivamente compensación por riesgo, sin embargo, la selección es lo que explica la relación empírica entre riesgo transitorio y la media de ingresos.

Key words: Premio por el Riesgo, Mercado de Trabajo, Industias, Ventajas Comparativas
JEL Classifications: D91 · J31 · J61.

1 Introduction

This paper is a quantitative study of the pricing of risk in the labor market. Specifically, we estimate the correlation between earnings risk and the level of earnings across industries and develop a theoretical framework to decompose that correlation into a compensating differential and a selection effect.

In the initial stages of their labor market history, workers sort themselves into careers that are often attached to a sector or an occupation. Someone who studies economics may, for example, consider entering the financial sector or working for the government as a policy economist as appealing career choices. The characteristics of working in either sector, as well as the worker's skill set, are the primary determinants of that choice. This paper focuses on one characteristic of employment that varies across industries: volatility in earnings. Being employed in sectors such as finance or business services is perceived to be riskier than being employed in social services or the public sector. If workers dislike risk, compensation for bearing that risk will translate into higher earnings for the economist working in finance compared to the policy economist working in the public sector.

The first goal of this paper is to closely examine this correlation: are industries featuring higher risk in earnings (both transitory and permanent) associated with higher earnings levels? Although two economists may appear to have identical skill sets (courses taken or how much computer programming they know), they may have differences in some unobserved ability that makes one of them more productive in the finance rather than in the government sector. In other words, the comparative advantages of workers may differ and they end up self-selecting into different industries based on those advantages. Through its equilibrium effect on earnings, the shape of the distribution of comparative advantages across the population partly determines the allocation of individuals across industries, affecting the estimated correlation between the variability and the level of earnings. The second goal of the paper is to estimate what fraction of the observed correlation is compensation for risk and how much of it is selection. To that end, we construct an equilibrium environment in which the two channels are explicitly modeled in order to contrast them with data. The estimated relationship between permanent risk and the level of earnings reflects compensation. However, temporary risk is not priced; the observed correlation is entirely due to selection.

The analysis of heterogeneity in ability levels of individuals or comparative advantage and its effect on career choice goes back to Roy's seminal work (Roy (1951)). We see our work as the first that integrates Roy's ideas into the analysis of career choice under uninsurable idiosyncratic labor earnings risk in general equilibrium. The heterogeneity in earnings risk we document, and its relation with the observed level of earnings and occupational choices, is central in the analysis of a wide range of policies considered in macroeconomics, public finance, and labor economics. Understanding what fraction of inequality observed early in life arises solely from career choices, is a necessary element in the design of policies targeting income redistribution. Moreover, our framework allows us to analyze the importance of unobserved abilities in shaping the career decisions of individuals and serves as a useful tool for contrasting the effect of policies directed at modifying initial conditions versus those aimed at providing insurance against shocks over their working life.

The paper has two distinct parts. In the first part, we employ the Survey of Income and Program Participation (SIPP) to estimate quarterly labor earnings risk across 21 industries of the US economy. Our definition of earnings risk is broad, encompassing unemployment spells, unexpected declines in hours, and decreases in wages. Both the definition of risk and the estimation methodology are based on literature for modeling earnings dynamics using panel data. We find substantial differences in the degree of labor earnings risk across industries. Workers in the financial or transportation industry experience large permanent shocks to earnings, while those working in social services are insulated from earnings variability. Working for the government entails low permanent risk but high transitory risk. Moreover, the evidence favors a positive correlation between mean earnings and earnings risk, once we control for other industry characteristics that affect the average level of earnings. The estimated coefficients imply that, when considering permanent shocks to earnings, the difference in average earnings between the riskiest and safest industries is around 10%. When shocks are transitory, moving from the safest to the riskiest industry implies an increase in mean earnings of 8%.

It is tempting to interpret the estimated correlation as a compensating differential for risk in the labor market. However, the sorting of individuals into the different sectors of the economy is endogenous: their sectoral choice depends on the risk they face and their sector-specific abilities. From reduced-form estimates it is not possible to unravel the two channels, of which fixed-effects estimates from individuals' earnings regressions are a convolution. As a result, the apparent risk premium may well be an artifact of our inability to control for self-selection based on the unobservable characteristics of individuals. To understand what part of the earnings differential is compensation for risk and which part is due to selection, the second part of the paper presents an overlapping generations model.

In our environment, risk-averse individuals, in addition to making a standard consumption versus savings choice, choose an industry in which to supply labor services. Some industries are riskier than others and, everything else equal, they are less attractive. Individuals are ex ante heterogeneous since each of them has a sector-specific skill or comparative advantage. In the spirit of the original model in Roy (1951), an individual can be very productive in the Finance sector but not so productive in Agriculture. In the absence of these fundamental differences in the distribution of abilities, when facing a high volatility of earnings in some industries individuals prefer to seek safer alternatives, supplying more labor to low-risk sectors, and hence depressing wages. In equilibrium, the nature of the earnings distributions across industries is shaped by the two different channels: on the one hand, the aversion of workers to supply labor to risky industries and on the other hand the distribution of abilities that determine their comparative advantage.

In the model, the relative level of risk across industries is given by the variances of persistent and transitory shocks estimated in the first part of the paper. In addition, our calibrated economy matches the share of labor across different sectors of the US economy taken from national accounts. We also parameterize the distribution of abilities so that the model delivers the mean and standard deviations of the cross-sectional distributions of earnings observed in the data. As a result of the sorting of workers, a natural distribution of mean earnings and industry risk arises. Interestingly, the model predicts a distribution of workers into sectors that closely resembles the one observed in the US data.

Viewed through the lens of the model, the positive relationship between the variance of both the permanent and transitory shocks to earnings and the average level of earnings are a convolution of two forces: the compensation for risk and the compensation for sector-specific skills. Therefore, in order to break down the effect of these two forces into the observed differences in mean earnings we proceed to perform a counterfactual exercise in which we shut down individuals' differences in ability or comparative advantage. In other words, we consider the individuals as ex-ante homogeneous. In this counterfactual world only the differences in the volatility of earnings across sectors shape the individuals' sectoral choice. With reasonable levels of risk aversion, the model over-predicts the positive correlation between mean earnings and permanent risk, i.e. a risk premium that is higher than in the data. On the contrary, it predicts a temporary risk premium that is virtually zero. Therefore, according to this result the strong association between the variance of transitory shocks and mean earnings observed in the data which, in light of the reduced-form model can be interpreted as a pure risk premium, arises entirely from selection. A large fraction of individuals possesses skills which increase productivity in industries with relatively large transitory shocks. Hence, despite their aversion to risk, their comparative advantage leads them to work in high (temporary) risk industries.

Related Literature To our knowledge, the first attempt to empirically analyze the link between the variability of income and mean earnings was the seminal work of Kuznets and Friedman (1939) in their classic study of income of professionals and more recently, Abowd and Ashenfelter (1981), Feinberg (1981), Leigh (1983), and Carroll and Samwick (1997).

The first three references analyse empirically the relationship between risk and earnings but lead to contradicting conclusions as the small datasets employed are less ideal than the SIPP. Moreover, they interpret their empirical results as proof (or lack thereof) of the existence of a risk premium or compensating differential. The fourth reference, Carroll and Samwick (1997) tests the hypothesis that households whose members are employed in high-risk industries accumulate more precautionary wealth. Our work contributes to a growing literature that develops quantitative models of occupational choice and income dynamics. An important paper is Kambourov and Manovskii (2009) who study the interplay between occupational mobility and wage inequality. Even though we focus on industries instead of occupations (in Appendix D we discuss the relationship between industry and occupations in our data set) and we abstract from mobility, our work can be seen as complementary to theirs. We bring to light a source of wage inequality that is still intimately related to the occupational-industry choice of individuals. More recently, in a work contemporaneous to ours, Dillon (2012) finds a positive relationship between the expected value and variance of lifetime earnings. Besides the different methodology and data set used by this author, our framework incorporates the sectoral decision of workers in a general equilibrium model as well as explicitly models the interplay between unobserved abilities and income uncertainty, which are absent in her work. Nevertheless, this author uses a richer econometric model and, more importantly, her results complements and confirms our main empirical finding.

One important contribution of our paper is to measure idiosyncratic labor market risk by industry and its macroeconomic implications in a general equilibrium framework. On the measurement side, we build on papers such as Storesletten, Telmer, and Yaron (2004b), Guvenen (2009), and Low, Meghir, and Pistaferri (2010), but we extend this literature by explicitly considering different industries. On the modeling side, our work belongs to the extensive quantitative macroeconomics literature with heterogeneous agents and incomplete markets that was initiated by Bewley (1977), Huggett (1993), and Aiyagari (1994). More recent contributions include Storesletten, Telmer, and Yaron (2004a), Heathcote, Storesletten, and Violante (2008), and Heathcote, Storesletten, and Violante (2009).

Finally, as mentioned, our framework incorporates workers comparative advantage and its effect on occupational choice and so it is closely related to (Roy (1951)). The empirical content of the original Roy model is studied in Heckman and Honore (1990) and Buera (2006). Roy's ideas are also adapted in modern dynamic discrete choice models to analyze the sources of income inequality, firstly in an important paper, Keane and Wolpin (1997) and, recently in Hoffmann (2010). However, we see our work as being the first that integrates Roy's ideas into the analysis of career choice under uninsurable idiosyncratic labor earnings risk in general equilibrium. In this line, we see our framework as a useful tool to be applied for future work interested in incorporating workers' comparative advantage into the analysis of earnings dynamics and of wage inequality.

2 The Story in a Simple Static Model

This section previews the main forces at work in the quantitative model presented below. Our artificial economy is populated by a mass of risk-averse individuals of total measure equal to one. Time is discrete and individuals live for only 1 period. Each individual is endowed with one unit of time each period that is supplied inelastically in a competitive labor market. There is a representative firm that produces a consumption good according to the following constant return to scale technology:

$$Y = (L^1)^{\phi}(L^2)^{1-\phi}, \tag{1}$$

where Y is output, L^1 and L^2 represent the two types of labor inputs required to produce output and, ϕ is the share of type-1 labor.

Individuals choose to supply one of the two labor types. If an individual chooses to supply type-1 labor, she earns the wage rate for that type, w^1. Alternatively, if she supplies type-2 labor she earns $w^2 z\gamma$ where w^2 is the wage rate for labor of type 2, $z > 0$ indicates a type-2 labor specific skill or ability distributed as $G(z, \theta)$ and, γ is a shock to labor earnings such that $\gamma = 1$ with probability p and $\gamma = \gamma_L < 1$ with probability $(1 - p)$.

Note that the ability when supplying type-2 labor, enters directly the productivity and hence earnings of an individual, determining an individual's comparative advantage for supplying that type. Also notice that workers supplying type-2 labor experience higher variability in earnings relative to those supplying type-1 labor, so if they are risk averse, everything else equal 2 looks less attractive than 1.

If a worker chooses to supply type-1 labor, her consumption is just the wage rate, i.e.

$$c^1 = w^1 \tag{2}$$

If instead chooses to supply type 2 labor then

$$c^2 = w^2 z\gamma \tag{3}$$

The individual chooses to supply labor of type 1 or 2 depending on the alternative that renders the highest utility. Therefore,

$$j^* = argmax\left\{V_1, V_2(z)\right\} \tag{4}$$

with $j^* \in \{1, 2\}$.

where

$$V_1 = u\left(w^1\right), \tag{5}$$

$$V_2(z) = p\left[u(w^2 z)\right] + (1-p)\left[u(w^2 z \gamma_L)\right], \tag{6}$$

being u the utility function with $u_c > 0$ and $u_{cc} < 0$.

The aggregate level for both types of labor results from individuals's choices:

$$L_1 = G(z^\star, \theta) \tag{7}$$

$$L_2 = E_\gamma \int_{z^\star}^{\infty} z dG(z, \theta) \tag{8}$$

where $z^\star$ is the level of z such that if $z > z^\star$ individuals choose labor type 2 and if $z \leq z^\star$ the choose labor type 1.

Therefore mean earnings for each type of labor can be expressed as

$$e_1 = w^1 \tag{9}$$

and

$$e_2 = \frac{w^2 \int_{z^\star}^{\infty} z dG(z, \theta)}{1 - G(z^\star, \theta)} \tag{10}$$

The representative competitive firm maximizes profits so

$$w^1 = MPL_1 = \phi L_1^{\phi-1} L_2^{1-\phi} \tag{11}$$

and

$$w^2 = MPL_2 = (1-\phi) L_1^{\phi} L_2^{-\phi}. \tag{12}$$

Assume that the utility function is of the logarithmic type. Then,

$$V_1 = log\left[w^1\right]. \tag{13}$$

and

$$V_2 = p\left[log(w^2 z)\right] + (1-p)\left[log(w^2 z\gamma_L)\right]. \tag{14}$$

Substituting (11) and (12) into (13) and (14), respectively, we have that

$$V_1 = log\left[\phi L_1^{\phi-1} L_2^{1-\phi}\right]. \tag{15}$$

and

$$V_2 = p\left[log((1-\phi)L_1^{\phi} L_2^{-\phi} z)\right] + (1-p)\left[log((1-\phi)L_1^{\phi} L_2^{-\phi} z\gamma_L)\right]. \tag{16}$$

Given (7), (8), (15) and (16) and the individuals labor type decision problem (4), the cut off level of skills, $z^\star$ is the solution of a function $Z(\phi, \theta)$ or, in the other words, the crossing point of (15) and (16). It is easy to prove that V_2 is a continuous monotone increasing function in z (with $\partial V_2/\partial z = 1/z$) and given that V_1 is constant in z then there is a single crossing point between V_1 and V_2 that gives a unique solution $z^\star$.

The analysis in this section is summarized in Figures 1 through 5, which depict a simple numerical example illustrating the mechanisms at work in the model.[1]

Individuals utility is represented in Figure 1 by the curves V_1 and V_2 as a function of z. As noted above, V_1 is independent of z and so it is a constant as it is shown in the figure. On the contrary, V_2 is strictly increasing in z and the crossing point of the two determines the value $z^\star$. Individuals whose ability levels are below $z^\star$ choose to supply type-1 labor and those with ability levels above the threshold, supply type-2 labor. This ability threshold determines the mass of both types of individuals, and hence the equilibrium values L^1 and L^2.

The first feature of the model we want to highlight is the positive correlation between mean earnings and the variance of the labor earnings shock. Figure 2 shows the ratio between mean earnings of labor of type 2 and 1.[2] For higher levels of the variance of the shock to earnings (x-axis) the model predicts higher mean earnings of labor 2 with respect to labor 1 (y-axis). In other words, for individuals that dislike risk a premium of labor earnings must be offered to compensate them for bearing that risk. What is the underlying mechanism that generates this feature in the model? Figure 3 shows the changes in the equilibrium value of $z^\star$ for different values of the variance of the shock to labor earnings. As can be easily noted, the higher the variance of the shock, the fewer the workers who choose to supply type-2 labor. In the figure, the lower number corresponds to the change in thresholds $z_1^\star$ to $z_2^\star$ and then $z_3^\star$. Given the decreasing returns to each type of labor in the production technology, the fewer workers choosing to supply type-2

[1] It is assumed that $G(z, \theta)$ is distributed $\Gamma(\kappa, \theta)$ $(\theta = (\kappa, \theta))$ with $\gamma = \theta = 2$, $\phi = 0.5$ and $p = 0.5$

[2] In the numerical example we change the variance of the shock to earnings by changing γ_L.

labor, the higher their mean earnings relative to those who supply labor of type 1.[3]

In order to illustrate the effect of comparative advantages or ability levels in the pricing of risk in the labor market, we plot the ratio of mean earnings as a function of the variance of the shock to earnings (as in Figure 2) but for different values of the mean of ability, $E(z)$. In particular, we assume $E_1(z) < E_2(z) < .. < E_5(z)$.[4] Those different values for $E(z)$ yield different curves for the relationship between the earnings premium and the level of risk. As the expectation increases, those curves shift to the right, as is apparent in Figure 4.

As is documented in our empirical analysis below, the data show a positive correlation between mean earnings and the variance of the shock to earnings across industries. The challenge in interpreting that correlation is that abilities are not observed. Therefore a structural framework is needed to assess whether that correlation reflects compensation for risk. Based on the analysis of our simple static model we argue that the sign of the correlation is the result of two forces operating at the same time: the compensation for risk or risk premium and the compensation for ability or comparative advantage. Suppose that different pairs of $\{\sigma_\gamma^2, E(z)\}$ represent different economies or industries (as in our empirical analysis) and that the mean of ability levels are positively correlated with the variance of the shock to labor earnings (i.e. the higher the $E(z)$ for an industry the higher is σ_γ^2). Then, it may well be the case that the ratio of mean earnings and variance of the shock are not correlated or even negatively correlated. This case is shown by the red line in Figure 5. That is, there exists compensation for risk even though the data shows a negative correlation between mean earnings and the volatility of earnings. Therefore, to make sense of that empirical correlation we need a quantitative structural model that can be confronted with data. We leave that task to Section 4.

3 Labor Market Risk and Mean Earnings

In this section, after briefly describing our data set, we document that risk and return in earnings are positively correlated across industries. We do this in two steps. First, we estimate the labor earning processes and properties of the shocks that workers of different sectors face in their work lives. Second, we characterize and estimate the empirical relation between mean earnings and the dispersion of earning shocks across sectors.

Our definition of labor earnings is rather broad (but consistent with previous studies). Besides the obvious variability in wage rates, we also consider changes in earnings due

[3] Here, the magnitude of the changes in the equilibrium value of $z^\star$ depend upon the value of ϕ.
[4] In our numerical example this is achieved by changing the parameter κ.

to changes in the amount of hours worked or changes in employment status.[5] As we make clear below, those changes which may be predicted based on information about individuals are not included in our measure of risk. For instance, if on average women who are between 25 and 30 years old begin working part-time after having been full-time employees, this decrease in the amount of hours worked, and the resulting earnings decline, is not considered risk by our methodology. We focus on individuals who never change industries, as this is most consistent with the quantitative framework we use below.

3.1 The SIPP

To explore whether the level of average earnings and the degree of unforeseen variability in those earnings are positively related, we turn to data. Ideally, to get an accurate answer to that question one would hope for a long high-frequency large panel of individual labor earnings with characteristics describing both the employee and the employer. The richer that data set, the easier would be to separate risk from other features that could affect average earnings. For the United States, the Survey of Income and Program Participation (SIPP) is the best approximation to that ideal data set. It is constructed by the U.S. Census Bureau and it takes the form a series of continuous panels which follow a national sample of households. The first panel began in 1983 but these earlier panels had a short duration. Starting in 1996 the Census Bureau began constructing longer panels with a larger number of households (more than 30,000 although the actual size varies) and those panels are the ones on which we focus on.

The SIPP conducts quarterly interviews which ask interviewees (individuals) to provide information at the monthly frequency on variables such as labor earnings, demographic characteristics, occupation, etc. It follows individuals for only 16 quarters, and this short duration prevents us from having entire life-cycle profiles of earnings. SIPP variables variables are collected for at most two jobs, but the survey also asks which of those is the primary job for the individual. In Appendix A we describe step-by-step our choice of the sample of individuals on which to perform the analysis described in this section. In brief, we focus on the reported primary jobs of married individuals between 22 and 66 years old and we eliminate those who are self employed, simultaneously report missing earnings but positive hours worked, report being out of the labor force, and do not report complete samples. In addition, we define earnings to include unemployment insurance if an individual reports zero hours worked and reports being unemployed.

Besides the good quality of earnings data in the SIPP, as analyzed in validation studies

[5]We do not consider individuals who move in and out of the labor force, but we do consider employment to unemployment transitions and vice versa.

comparing it to administrative data (see Abowd and Stinson (2011) and Gottschalk and Huynh (2006)), relative to other longitudinal panels such as the Panel Study of Income Dynamics (PSID) and the National Longitudinal Survey of Youth (NLSY97 and NLSY79), the advantages of the SIPP are mainly two. The first is the number of respondents. It is considerably larger than the PSID, which surveys about 10,000 households, or the NLSYs, which interview between 9,000 and 13,000. The second advantage is the frequency of interviewing. The SIPP provides a wealth of information at the monthly frequency; the PSID interviews annually (biannually since 1997) and the NLSY97 is now interviewing biennially. It is fortunate that for many individuals in the United States being unemployed or suffering a decline in income is a short-lived experience (usually weeks or months). But given those are the risks on which this study focuses, that fact underscores the importance of having information at higher frequencies.

3.2 Labor Income Shocks

The first step in our analysis computes earnings variability at the individual level with a regression approach used extensively in the literature, for example, in Carroll and Samwick (1997). We proceed by estimating a fixed effect model for each industry j in our sample. Given a panel of N individuals for whom we measure earnings (and other variables) over a period of time T, we assume that (log) earnings for individual i in industry j at time t, y_{ijt}, can be modeled as,

$$y_{ijt} = \alpha_{ij} + \boldsymbol{\beta}_j \boldsymbol{X}_{ijt} + u_{ijt} \tag{17}$$

The vector $\boldsymbol{X}$ comprises several variables that help predict changes in the level of log-earnings. Specifically, we include age, sex, ethnicity, years of schooling, an occupational dummy, and a seasonal dummy.[6] We first assume that the error term u_{ijt} is distributed i.i.d. $N(0, \sigma^2_{j,u})$.

We estimate equation (17) by ordinary least squares for all individuals in a given industry. Repeating this procedure for all industries yields estimates $\{\hat{\alpha}_{ij}, \hat{\boldsymbol{\beta}}_j\}_{j=1}^{21}$ and $\widehat{\sigma^2_{j,u}}$. We present the estimates of the variances of the innovations for each industry in Table 1.

The median of the estimated variances is 0.066 which corresponds to the earnings volatility for those workers who work in the Education sector. The workers who face the least amount of uncertainty are, in order, those who work in Armed Forces, Agriculture

[6] An alternative interpretation of the seasonal dummy is a periodic change over time in the coefficient α_{ij}.

and Forestry, Social Services, Mining and Utilities.[7] Workers in Finance, Medical Services, Other Services, Transportation and Hospitals are the industries with the highest levels of income uncertainty. Note that, according to this notion of risk, the Finance sector is more than twice as risky as Agriculture and Forestry. Finally, we test the hypothesis that all the estimated variances are equal and we reject it with a with a p-value of virtually zero.[8]

3.3 Permanent and Transitory Shocks

We now enrich our empirical analysis by allowing the error term to be decomposed into a permanent component and a transitory component. The reason for distinguishing between the two types of shocks is that they affect the welfare of workers differently. Transitory shocks (e.g. the loss of an important customer for a consultant) are seldom a cause for concern; small levels of savings are usually enough for workers to weather that type of shock successfully. Permanent shocks are, by definition, longer-lasting and can be associated with, for instance, depreciation of job-specific human capital or permanent changes in the way an industry operates. Smoothing out the latter type of shocks through a buffer stock of savings is more difficult and permanent changes in consumption are often times required. As the impact on the welfare of individuals is different for the two types of risk, one would expect that the premium that workers demand for bearing them differs as well.

We follow Carroll and Samwick (1997) and Low, Meghir, and Pistaferri (2010), among others, by assuming that

$$u_{ijt} = \eta_{ijt} + \omega_{ijt}, \tag{18}$$

where η_{ijt},the transitory component, is distributed i.i.d. $N(0, \sigma^2_{j,\eta})$, and ω_{ijt}, the permanent component, is a random walk, i.e.

$$\omega_{ijt} = \omega_{ij,t-1} + \epsilon_{ijt} \tag{19}$$

with i.i.d. innovations ϵ_{ijt} that are distributed $N(0, \sigma^2_{j,\epsilon})$. By estimating equation (17) we obtain $\{\{\hat{u}_{ijt}\}_{i=1}^{N_j}\}_{t=1}^{T}$.

We estimate the variances of the permanent and transitory components by following the identification procedure proposed in Low, Meghir, and Pistaferri (2010).

Taking first-differences in equation (17) and given the process specified in (18), we have

[7] Regarding Armed Forces, even though it is a low earnings risk sector, could be considered risky using alternative metrics (e.g. injuries and death while in service).

[8] To test this hypotheses we use the Welch test.

$$\Delta y_{ijt} = \Delta\boldsymbol{\beta}_j \boldsymbol{X}_{ijt} + \Delta\eta_{ijt} + \epsilon_{ijt}. \tag{20}$$

Now define

$$g_{ijt} = \Delta(y_{ijt} - \boldsymbol{\beta}_j \boldsymbol{X}_{ijt}) = \Delta\eta_{ijt} + \epsilon_{ijt}. \tag{21}$$

To identify the parameters of interest, we compute,

$$E(g_{ijt}^2) = \sigma^2_{\epsilon_{ij}} + 2\sigma^2_{\eta_{ij}}, \tag{22}$$

and

$$E(g_{ijt}g_{ijt-1}) = -\sigma^2_{\eta_{ij}} \tag{23}$$

To estimate the variances of the two innovations, we proceed as follows. For an individual i in a given industry j, we estimate $\widehat{E(g_{ijt}^2)}$ and $E(\widehat{g_{ijt}g_{ijt-1}})$ by taking the sample moments. By solving the system, we then obtain $\widehat{\sigma^2_{\epsilon_j}}$ and $\widehat{\sigma^2_{\eta_j}}$ by taking averages across individuals of the estimated variances obtained for each individual.

Table 2 shows the estimated variances and Figures 6 and 7 show the magnitudes of the estimated variances for each industry sorted from the smallest to the largest of the permanent and transitory shock, respectively. The median of the estimated variances across industries are 0.0141 and 0.0037 for the permanent (column 2, Construction) and transitory shocks (column 4, Medical Services), respectively.

Regarding the variance of the permanent shock the group of relatively safe industries is comprised of the Armed Forces, Social Services, Utilities, Communication and Government. The most uncertain sectors are Finance, Transportation, Retail Trade, Education and Business Services. The risky sectors according to the variability of the temporary component are Mining, Agriculture and Forestry, Finance, Government and Other Services. On the other hand, the sectors with the lowest variance of transitory income shocks are Recreation and Entertainment, Armed Forces, Business Services, Personal Services and Construction. Without exception, the variance of the permanent component is higher than that of the transitory component by a factor of roughly three. Finally, we find interesting the intersection of both the permanent and transitory risk across sectors. To put it simply, Table 3 describes the distribution of sectors across these two dimensions, classifying them into risky or safe if they are above or below the median of the estimated variances of these shocks. According to this classification, there are five sectors that we can consider risky in terms of both type of shocks: Hospitals, Agriculture and Forestry, Medical Services, Finance and Retail Trade. On the contrary, there are four sectors with both type of variances below the median and so they can be considered as safe sectors: Armed Forces, Utilities, Personal Services and Recreation and Entertainment. In addition, there are sectors that are safe in terms of permanent shocks to labor earnings but for which temporary shocks are more severe or above the median, these are: Social Services, Communication, Government, Non Durable Goods Manufacturing, Other Services and Mining. Finally, the sectors for which the variance of the permanent shock is above the median but the variance of the temporary shock is below the median are: Construction, Wholesale Trade, Durable Goods Manufacturing, Business Services, Education and Transportation. Besides the rich characterization of the risk workers face in the labor market that this type of descriptive analysis brings to the table, it also shows the type of trade offs that individuals face when they decide the industry for which they offer their labor services. As specifically considered in our model, the insurance opportunities individuals have will allow them to smooth out shocks to labor income and at the same time shape their sectoral choice.

3.4 Industry Risk and Risk Premia

Having estimated measures of risk for our group of industries, we are now ready to test the hypothesis that, across industries, the level of risk and the average level of earnings are positively correlated. Our claim, however, should be understood to be a *ceteris paribus* claim. That is, everything else constant, a higher level of risk should be associated with a higher level of earnings. Of course, not everything else is constant across industries. Industries differ along many dimensions that may affect average earnings independently of their level of risk. This should lead one to suspect that the mix of workers or firms in a given industry are important determinants of its average level of earnings. From the econometric point of view, to account for this industry heterogeneity, we proceed in two ways. We first compute industry averages (that is, averages across individuals who work in a given industry) of variables we deem relevant in determining average earnings. More specifically, we establish the (conditional) sign of the relationship between average earnings and industry risk by estimating the following regression equation:

$$y = \gamma + \boldsymbol{\theta}\boldsymbol{Z} + \nu \tag{24}$$

where y is a vector whose j^{th} element is the average (log) earnings for individuals in industry j, and $\boldsymbol{Z}$ is a matrix of regressors. The j^{th} row of $\boldsymbol{Z}$ has six elements. the average age, the average age squared, and the average level of education of all individuals working in industry j, the fraction of females in industry j, and the industry j variance of income shocks estimated above (see Table 1). Since the number of industries in our sample is 21,

y is a column vector of dimension 21, and $\boldsymbol{Z}$ is of dimension 21×6.[9] Finally, γ is a vector of intercepts and ν a vector of residuals. We assume that the error term ν_j is distributed i.i.d. $N(0, \sigma_\nu^2)$.

Column 2 of Table 4 shows the results of estimating equation (24). It presents the estimated values for the coefficients and their probabilities of being less than zero computed by bootstrap. All the coefficients are significant and have the expected signs. Workers' age and education levels are positively related to mean earnings and females labor earnings are on average lower than males. Our focus is on the sign and magnitude of the coefficient associated with risk, σ_ϵ^2. The table shows that this coefficient points to a strong and positive association between uncertainty and earnings. More importantly, the probability that this coefficient is less than zero is 0.0002. Note that the value of the coefficient associated with uncertainty implies that increasing the variance from 4.9% to 9% (we go from Agriculture and Forestry to Finance), increases the mean level of earnings by 28%. According to this econometric model, this result would be consistent with the existence of a risk premium in the labor market. [10]

Alternatively, in order to document the relationship between our measure of industry risk and mean earnings we also estimate individual's earnings net of its main observed characteristics: age, the square of age, education and gender. We proceed by estimating a pooled regression by OLS that allows us to obtain estimates for the net log earnings of each individual in our sample. Specifically, we estimate the following regression

$$y_{ijt} = \gamma_0 + \boldsymbol{\gamma} \boldsymbol{X}_{ijt} + \lambda_{ijt} \tag{25}$$

where the vector of coefficients γ represents the effect of the observed characteristics (age, education, square of age, and gender). These coefficients are common in the cross-section and across time. When then obtain, for each individual and at a point in time, log earnings net of observed characteristics by computing $\tilde{y}_{ijt} = y_{ijt} - \hat{\boldsymbol{\gamma}} \boldsymbol{X}_{ijt}$. By averaging across time and across individuals in each industry we obtain the mean of the net earnings by industry. Specifically, for a given industry j, we compute

$$\tilde{y}_j = \frac{1}{N_j} \sum_{i=1}^{N_j} \tilde{y}_{ij}, \tag{26}$$

where

[9] We also estimate equation (24) when the matrix $\boldsymbol{Z}$ includes the variances of the temporary and permanent component of income shocks. In that case, the matrix $\boldsymbol{Z}$ has 7 columns because we do not include the vector of overall variances.

[10] We also estimate equation (24) by defining earnings in per hour terms, our results are robust to this specification. See Appendix C for details.

$$\tilde{y}_{ij} = \frac{1}{T}\sum_{t=1}^{T} \tilde{y}_{ijt}. \tag{27}$$

We now use this estimated values of net earnings to estimate the following regression equation:

$$\tilde{y}_j = \alpha_0 + \alpha_1 \sigma_j^2 + \nu_{\tilde{y},j} \tag{28}$$

Again, since the number of industries in our sample is 21, each variable of the regression has 21 observations. α_0 is an intercept, α_1 is the coefficient that represents the effect of our measure of risk into the mean of net earnings and $\nu_{\tilde{y}}$ is the residual which is assumed to be i.i.d. $N(0, \sigma^2_{\nu_{\tilde{y}}})$.

Column 3 of Table 4 shows the estimated values for the coefficients and the probabilities of that they are less than zero computed by bootstrap. Note that, somehow confirming the existence of a risk premium, the coefficient associated with risk, α_1, is positive with a probability of being less than zero equal to 0.0104. Note, however, that according to this approach, the value of the coefficient associated with uncertainty implies that increasing the variance from 4.9% to 9% (we go from Agriculture and Forestry to Finance), increases the mean level of net earnings by 7%.

As mentioned above, the decisions of workers could greatly differ upon the nature of the shock, so it is important to consider the decomposition of the process into a temporary and a permanent component.

For these reasons, we estimate equation (24) using as regressors the variances of the two components, permanent and transitory, instead of just one variance that reflects overall uncertainty. The second column of Table 5 presents the results.

All the coefficients are significant and have expected signs. Excepting the coefficients associated with the variances, their magnitudes are close to the ones found before. Turning now to the coefficients associated with uncertainty, we first observe that they are strongly positive and with probabilities of being less than zero of 0.18 and 0.16 for the permanent and transitory shocks, respectively. The estimated value for coefficient associated with the variance of the permanent shock to earnings is 9.3. Therefore, according to this result, going from the Social Service industry (the second safest) to Finance (the riskiest industry) implies an increase in mean earnings of 7%. Regarding the transitory shock, the value estimated for its associated coefficient is 20.3 and so, according to this result, moving from Recreation and Entertainment (the safest sector) to Mining (the riskiest sector) implies an increase of in mean earnings of 10%. As in the case with the total variance of earnings, we present the results by using our alternative specification to document the relationship between mean earnings and uncertainty. It is depicted in column 3 of Table 5. The estimation results point to a strong and positive relationship between mean earnings and the estimated variances, being the values of estimated coefficients for the permanent and transitory shocks to earnings are 6.9 and 16.6, respectively, with very low probabilities of being less than zero: 0.015 and 0.077, respectively. According to these results, considering the permanent shock to earnings, moving from Social Services to Finance implies an increase in mean earnings of 5%. If we look at the transitory shock to earnings, moving from Recreation and Entertainment to Mining implies a compensation in mean earnings of 8%.

The data and approach we use to link labor earnings and their uncertainty yield estimates which appear to be consistent with a compensating differential for risk in the labor market. But, as illustrated in Section 2 one ought to be cautious. The distribution of average earnings across individuals in an industry is an endogenous outcome resulting from individuals' decisions of where to supply their labor services. The level of earnings risk is certainly something individuals consider when making that choice. But their comparative advantage, in other words, their relatively higher productivity in a certain sector, consequence of a set of individual characteristics, plays a role as well. Some of that comparative advantage originates from being, for instance, a female or a college-grad, characteristics which we have accounted for to some degree. Much of the advantage, however, originates from unobserved characteristics which are, obviously, difficult to control for. To help us decompose how much of the estimated earnings risk *premium* is due to a compensating differential and how much to self-selection, the next section describes a quantitative framework in which comparative advantage and individuals' industry choice are explicitly taken into account.

4 The Model

Our artificial economy is populated by a mass of risk-averse individuals of total measure equal to one. Time is discrete and individuals live for S periods which correspond to their working lives. In other words, they are born into a labor market and never retire. Each individual is endowed with one unit of time each period that is supplied inelastically in the labor market. When an individual reaches time $S+1$ and dies, another age 0 individual replaces her, so the total population is constant. At the beginning of their lifetimes, individuals choose to work in one of J mutually exclusive job opportunities indexed by j, which we interpret as sectors or industries. At birth, prior to the industry choice, each individual draws a value for sector-specific skill or ability from a given distribution specified below. These skills enter directly the productivity and hence earnings of an individual and therefore determine an individual's comparative advantage for, say, working in Finance and not in Agriculture. As these skills are random draws, we are silent about their origin but they could loosely be interpreted as innate abilities or human capital acquired before entering the labor market. Finally, the values for the sector-specific skills do neither grow nor decrease over an individual's lifetime.

In addition, once working for an industry (from which they cannot move), individuals are subject to idiosyncratic shocks to their labor income. The process driving those shocks differs from industry to industry. In particular, workers in some industries experience a higher variability of earnings relative to workers in other industries. If workers are risk-averse, riskier industries look less attractive than safer industries.

When an individual is born in period 0 (i.e. when she enters the labor market), her problem is to choose one of the J mutually exclusive career alternatives in order to maximize the expected discounted value of her life-time utility:

$$\mathbb{E}_0 \left\{ \left[\sum_{s=1}^{S} \beta^{s-1} \sum_{j} \mathbf{1}_j u(c_{s,j}), \right] \middle| \Omega_{i,0} \right\},$$

where $\mathbf{1}_j$ is an indicator function with value $\mathbf{1}_j = 1$ if the individual chooses to work in industry j and 0 otherwise. The function $u(c_{s,j})$ denotes the individual's per-period utility derived from choosing $j \in J$ and consuming $c_{s,j}$; we assume $u_c > 0$ and $u_{cc} < 0$. The only source of uncertainty are shocks to labor earnings and we describe those in detail below. For now it suffices to say that expectations in (4) are taken with respect to the distribution of those shocks. The vector $\Omega_{i,0}$ represents the information set of an individual i at time 0 and it is formally the vector

$$\Omega_{i,0} = \{\theta_{i,1}, \dots, \theta_{i,J}\}$$

where the logarithm of each value $\theta_{i,j}$ is drawn from an industry-specific distribution $N(\mu_{\theta_j}, \sigma^2_{\theta_j})$. Each period, by inelastically supplying one unit of time to sector j, each individual receives labor earnings, $w_j \theta_{i,j} e^{\nu_{i,j}}$, comprised of a sector specific competitive wage rate (w_j), individual-specific sectoral pre-labor-market skills $(\theta_{i,j})$ and, an individual-specific but time-varying labor productivity shock $(\nu_{i,j})$. Once the individual makes her sectoral choice, only the θ corresponding to the chosen industry affects her lifetime labor earnings.

For an individual of age s, the time-varying component of earnings is the addition of two orthogonal stochastic components,

$$\nu_{s,j} = \eta_{s,j} + \omega_{s,j}, \tag{29}$$

where η_j is an i.i.d. transitory shock to log earnings distributed as $N(-\frac{1}{2}\frac{1}{\sigma^2_{j,\eta}}, \sigma^2_{j,\eta})$ and $\omega_{s+1,j}$ is the permanent component that follows a random walk: $\omega_{s+1,j} = \omega_{s,j} + \epsilon_{s,j}$ with ϵ_j being $N(-\frac{1}{2}\frac{1}{\sigma^2_{j,\epsilon}}, \sigma^2_{j,\epsilon})$ i.i.d innovations.[11] By subscripting the variance by j, we make clear that the nature of the shock process is industry-specific. Despite the inability of consumers to change industry in midlife, we allow them to partially insure against labor income shocks by saving in a one period risk-free non-contingent bond with an exogenous interest rate equal to r.

Individual's Decision Problem Suppose an individual has chosen an industry in which to supply labor and begun her working life. Every period, optimization for this individual entails choosing how much to consume and the amount of savings or quantity of one-period bonds to purchase.[12] The variables relevant to these decisions are the level of wealth (b), the age of the individual (s), and the following components of income: the time-varying component (ω and η) and the ability level for the chosen industry (θ_j). Thus the vector of individual state variables can be denoted as $x = (b, \omega, \eta, s, \theta_j)$, where j is the chosen industry. Denote by $\Psi_j(x)$ the industry j workers' distribution across assets, age, income, and abilities.[13] It is an aggregate state variable since it determines the wage rate in industry j. Only the marginal distribution of age is identical across all industries. For convenience denote by $\mathcal{S} = \mathcal{S}_B \times \mathcal{S}_{E_\eta} \times \mathcal{S}_{E_\omega} \times \mathcal{S}_\theta \bigcup \{1, \ldots, S\}$ the state space of the vector of state variables x.[14] It is convenient to write the problem recursively, and we denote the remaining lifetime utility for an age-$\bar{s}$ individual working in industry j by the $V_j(x|s = \bar{s})$. It is defined by,

$$V_j(x|s = \bar{s}, \Psi_j) = \max_{c,b'} \left\{u(c) + \beta E V_j(x'|s = \bar{s} + 1, \Psi'_j)\right\} \quad \text{if} \quad 1 \leq s \leq S \quad \text{and} \quad 0 \quad \text{o/w},$$

subject to,

$$c + b' = w_j(\Psi_j)\theta_j e^\eta e^\omega + b(1 + r) \tag{30}$$

[11] In the quantitative application we approximate the random walk by a highly persistent process. It is close to a unit root but stationary nevertheless. See Appendix B for details.

[12] Our model is set apart from others in the literature in the optimal choice of an industry and its general-equilibrium implications. Once the individual has chosen an industry, the optimization problem of the consumer is essentially identical to many examples in a literature analyzing heterogeneous agents economies. The only departure is that we allow two different shocks with different statistical properties. This departure allows us to analyze the impact of transitory and permanent risk on industry choice.

[13] The distribution is subscripted by j because workers, facing different income shocks and self-selecting into industries based on different abilities levels, will choose different levels of assets.

[14] In general, the joint state space should have a subscript j. In our particular model, the borrowing constraint and longevity are identical across industries. Income innovations and abilities are all real numbers. Hence we can omit the subscript j.

and,

$$b \geq \underline{b}, \;\; b_0 = 0, \;\; b_{S+1} \geq 0 \tag{31}$$

We follow relatively standard notation when we denote by x' the values of x one period ahead. Equation (30) is a standard flow budget constraint that equates consumption plus savings to total earnings from capital holdings $b(1+r)$, and earnings from supplied labor $w_j(\Psi_j)\theta_j e^{\nu}$. In addition to this budget constraint, individuals face a borrowing constraint that restricts the lower bound on asset holdings. Also, individuals are born with zero wealth ($b_0 = 0$) and they face a non-negativity constraint in their savings at the time of death ($b_{S+1} \geq 0$).

At birth, the individual chooses from a set of J industries the one that yields the highest utility.

$$j^* = argmax \left\{ W_1, \ldots, W_J \right\} \tag{32}$$

where W_{j^*} for an individual i is defined as

$$W_{j^*} = \mathbb{E}_0 \left\{ V_{j^*}(x|s=1)|\Omega_{i,0} \right\}. \tag{33}$$

When choosing an industry, $\Omega_{i,0}$ - the vector of abilities drawn at birth - is in a person's information set, thus appearing to the right of the conditioning sign. The individual knows as well the statistical properties of shocks that she will experience in each industry. As a result, and although not explicitly written, it should be understood that the expectation is taken with respect to a different distribution if the worker computes W_j for $j \neq j^*$. The choice in (32) induces an endogenous distribution of workers across industries. Let μ_j denote the mass of workers in industry j then, $\sum_{k=1}^{J} \mu_k = 1$.

Firms One can picture our model economy as a small open economy containing a set of islands with each of the islands representing an industry. In each industry, a consumption good is produced according to the following industry-level technology:

$$Y_j = N_j^{\alpha_j}, \tag{34}$$

where Y_j is the output of sector j, N_j represents the labor input of that sector measured in efficiency units,[15] and α is the share of labor in output (with $\alpha < 1$). Firms are owned by foreigners who operate it, pay wages, and enjoy profits. We do not consider any kind of inter-industry trade in goods, so the reader can assume that goods produced across

[15]The measure of efficiency takes into account both the time-varying productivity component and the industry-specific abilities.

islands are identical. [16].

Equilibrium We can now define a stationary competitive equilibrium which consists of a set of industry wages $\{w_j\}_{j=1}^J$, industry populations (or masses) $\{\mu_j\}_{j=1}^J$, industry-specific distributions $\{\Psi_j(x)\}_{j=1}^J$, industry-level efficiency-weighted employment levels $\{N_j\}_{j=1}^J$, and industry-specific decision rules $\left\{b_j'(x), c_j(x)\right\}_{j=1}^J$ and associated value functions $\{V_j(x)\}_{j=1}^J$, which satisfy the following conditions:

1. Given wages, the industry-specific decision rules $\left\{b_j'(x), c_j(x)\right\}_{j=1}^J$ solve the optimization problem (4) yielding value functions $\{V_j(x)\}_{j=1}^J$.

2. The set of industry-specific populations $\{\mu_j\}_{j=1}^J$ and the distributions of abilities across industries are consistent with the optimal industry choice (32). For any given industry j, its population satisfies $\mu_j = Prob(W_j > W_{-j})$ where we define the vector W_{-j} to be equal to $\{W_1, \ldots, W_{j-1}, W_{j+1}, \ldots, W_J\}$. The cumulative distribution of θ_j in a given industry j is defined by,

$$G_j(\theta_{0,j}) = \frac{\int_{\Theta_{-j}} \int_{\{\theta_j \in \Theta_j : \theta_j < \theta_{0,j}\}} \chi_{\{\theta_j : W_j > W_{-j} | \theta_{-j}\}} dF(\theta_j) dF(\theta_{-j})}{\int_{\Theta_{-j}} \int_{\Theta_j} \chi_{\{\theta_j : W_j > W_{-j} | \theta_{-j}\}} dF(\theta_j) dF(\theta_{-j})} = \int_S \chi_{\{\theta_j \leq \theta_{0,j}\}} d\Psi_j(x)$$

where Θ_j is the support of θ_j and Θ_{-j} is the support of θ_{-j} and $\chi_{\{\theta_j : W_j > W_{-j}\}}$ is and indicator function that takes the value 1 when an individual with ability θ_j chooses industry j. Finally, $F(\theta_j)$ is the c.d.f of θ_j before sorting of agents.

3. Wages in industry j are equal to the marginal product of a marginal unit of average efficiency in that industry:

$$w_j = \alpha_j N_j^{\alpha_j - 1},$$

where the industry-level measures of employment are defined as $N_j = \mu_j \int_{\mathcal{S}} \theta_j e^{\eta} e^{\omega} d\Psi_j(x)$.

4. For an individual in an industry j, the decision rules $b_j'(x)$ and $c_j(x)$ solve the individuals' dynamic problem (4), and $V_j(x)$ is the associated value function.

5. In a given industry j, $\Psi_j(x)$ is the stationary distribution associated with the transition function implied by the optimal decision rule $b_j'(x)$ and the law of motion for the exogenous shocks.

[16] Alternatively, one can picture J different goods and assume that an individual working in industry j obtains utility from consuming the good produced in that industry only, and not those from other islands

6. At the industry level, the following resource constraint is satisfied:

$$w_jN_j = \int_{\mathcal{S}} \{c_j(x) + b_j'(x) - b_j(x)(1+r)\} d\Psi_j(x)$$

5 Quantitative Analysis

This section presents the quantitative analysis. For this purpose, we use the theoretical model developed in the previous section which is computed and calibrated to mimic the US economy. [17] Besides the standard complexities associated with computing standard life cycle economies there is another layer of difficulty in this particular model that has to do with the existence of the pre-labor market skills or abilities distributions. The main reason has to do with computing and comparing value functions for each possible combination of the abilities draws for each simulated individual that lives in our model economy. Even though the presence of these variables dramatically enrich our analysis, due to this computational difficulty we restrict our quantitative analysis to 4 industries of the US economy: Agriculture, Manufacturing, Services and Public Sector, which result from aggregating the 21 industries detailed above. In Table 6 we present mean of net earnings the variance of the permanent and transitory shocks for these 4 aggregate sectors. Note also that even though we have aggregated the 21 industries, the strong and positive relation between mean earnings and the variances of the permanent and transitory shock is preserved: if we regress the mean net earnings on a constant and both variances we get that the coefficient associated with the variance of the permanent shock is 8.5 and with the variance of the transitory shock is 8.4. This implies that, considering the permanent shock to earnings, moving from the Public Sector (the safest) to Manufacturing (the riskiest) implies an increase in net earnings of 3.5%. If we consider the transitory shock to earnings, moving from the Public Sector (which is again the safest) to Agriculture (the riskiest) implies an increase in mean net earnings of 2%. We now turn to parameterize the model economy.

5.1 Parameter Values

We start by setting the model period equal to a quarter, and the total lifetime for an individual to be 120 periods. These two values correspond to 30-year employment histories. We exogenously set the annual interest rate to be 5%. In our benchmark case we start by setting $\underline{b} \geq 0$ and pick β to be 0.957 so that we match an aggregate wealth to income ratio of 3. We restrict preferences to be of the constant relative risk aversion

[17] The procedure followed to compute the model is presented in detail in Appendix B.

class with coefficient or risk aversion equal to 2. In addition, we need to assign values for the parameters that govern returns to scale at the industry level, α_j's. These parameters represent the labor's share of total revenue in each of the industries and, following Hopenhayn and Rogerson (1993) which use the same decreasing return to scale technology, we use National Accounts data to find values for them. Specifically, we use the Compensation of Employees and GDP at the industry level from the National Income and Product Accounts for the period 1990 through 2009 to set the labor share of Agriculture equal to 0.30, of Manufacturing equal to 0.63, Services equal to 0.51 and Public Sector equal to 0.85.

One of the driving forces of a non-degenerate wage distribution across industries is an industry-specific level of risk. As a measure of this risk, we use the estimates for the variances of the two components of income we estimate from SIPP in Section 3 and we aggregate to the 4 industries we focus in this section. Hence, we set J, the total number of industries, to be 4 and we feed the model with the estimated values of the variances of both the permanent and transitory shocks depicted in the fourth and fifth column of Table 6.

Finally, it still remains to parameterize the distributions of pre-labor market skills or abilities, i.e. to find values for 8 parameters: $\{\mu_{j,\theta}, \sigma^2_{j,\theta}\}_{j=1}^{j=4}$. For this purpose we pick values for these parameters so that the model delivers the mean and standard deviation of the net earnings for each of the 4 industries (column 2 and 3 of Table 6). The use of net earnings is justified by the fact that in our model economy all individuals are equal in terms of sex and education, and there is no age-specific productivity (i.e. all the observables we control for in equation (25)). The resulting parameter values are shown in Table 7.

5.2 Results

The experiment consists of solving the model for the set of parameter values just described. Since the parameters of the distributions of the pre-market labor skills are picked so that the model exactly replicates the mean of net earnings of each of the 4 industries, it transpires that the model also replicates the relationship between the mean earnings and the variances of the transitory and permanent shocks to earnings. Therefore, this empirical relationship cannot be used as an independent test for the model. Interestingly, the model has testable implications with regard to the sorting of workers into the 4 sectors of the economy. Specifically, the model predicts the mass of individuals that work in each of the sectors in equilibrium. Table 8 shows the model predictions and their data counterpart. The correlation between the share of individuals in each sector in the model and in the data is 0.92 being particularly good the model predictions regarding the mass of workers that work in Agriculture and Services. In addition, the model has predictions regarding the wealth to income ratio in each of the four sectors of the economy, Table 9 shows the model predictions. As expected the amount of wealth accumulated in each of the sector is positively correlated with the riskiness of earnings.

As it was highlighted above, the documented strong and positive relationship between mean earnings and their variances, can be interpreted as evidence in favor of a existence of a pure risk premium in the US labor market. However, the presence of individual specific pre-labor market skills or individual comparative advantages affect the sorting of individuals into the sectors of the economy. This is unobserved for the econometrician and so it can greatly change the interpretation of the results since we could be mistakenly assigning all of the observed differences in mean earnings to compensation for risk while in fact they are compensation for the skills of the individuals. Fortunately, our theory is rich enough so that we can proceed to perform a counterfactual experiment in which we shut down all the differences across individuals and across sectors in the pre-labor market skills. This counterfactual exercise could be interpreted as a way to decompose the observed relationship between mean earnings and the variance of shocks into pure compensation for risk and compensation for unobserved skills. Specifically, we solve the model again but instead of picking the parameters of the skills distributions to match the moments of the net earnings distributions, we just endow all individuals with the same skill level for each of the 4 industries that compose our economy.

By doing this our model economy dramatically changes its properties since the individual sectoral choice is only affected by the cross-sectoral differences in the volatility of earnings. The individuals are now ex-ante homogenous but still subject to idiosyncratic shocks to their labor income; hence they are *ex-post* heterogeneous. It is still the case that workers in some industries experience a higher variability of earnings relative to workers in other industries. For our risk-averse workers, riskier industries look less attractive than safer industries. Everything else equal, all workers would concentrate in the safest industry, with all but one industry having no workers and hence no output. In our environment, this can not be an equilibrium because industry level technologies display decreasing returns to scale. As a result, the more workers populate an industry the lower the wages, and vice-versa. The resulting equilibrium features relatively safe industries with low wages and a large mass of workers. Riskier industries display the opposite characteristics.

Table 10 shows the share of workers predicted by the model in this counterfactual exercise. Note that the share of workers in each industry is at odds with the data (correlation of -0.24) being the Public Sector, the safest sector, the one that is most preferred by workers. Also note that in the case of these four industries, since there is not much variation in the variance of the permanent shock to earnings (column 4 of Table 6) the share of workers do not vary much: 0.16 in Agriculture, 0.22 in Manufacturing and 0.21 in Services. Therefore, it is clear the influence of comparative advantages in shaping the individual sectoral choice in order the model is in line with the data. One way to illustrate this effect is by comparing the mean of pre-labor market skills before and after the sorting takes place in our model economy. This is presented in Table 11.

Given our identifying assumption that these pre-labor market skills are independent across sectors the ratio of means (column 2) are going to be greater than one for all the sectors that compose our model economy. In other words, only the most capable individuals for each sector are going to be self-selected in that sector. Nevertheless, this selection effect could be stronger in some sectors compared to others. According to our results, the strongest selection effect takes place in Agriculture followed by Manufacturing and the Public sector and, being relatively mild in the case of Services.

We move now to analyze the model predictions regarding the correlation between mean earnings and its volatility in this counterfactual experiment. Table 12 shows the correlation between the mean earnings predicted by the model and the variance of the permanent and transitory shock in both the benchmark case (column 2) and in the counterfactual experiment (column 3). Two important results emerge from this table: i) in the counterfactual exercise the coefficient associated with the permanent shock is 15.1 which points to a strong and positive association between mean earnings and the variance of the permanent shock to earnings as we observe in the benchmark case and in the data and, moreover, this result implies that the increase in mean net earnings from moving from the Public Sector to Manufacturing would be around 6.2% while in the data it is actually 3.5%. ii) Regarding the transitory shock, the coefficient corresponding to the transitory shock reduces to only 0.9 whereas in the benchmark case this coefficient is 8.4. By saving in one period bonds the workers that live in our model economy can perfectly smooth transitory shock to labor earnings and so they do not need to be compensated for bearing that type of risk in the labor market. For this reason and, in light of our model, what appears to be compensation for the risk associated with transitory shocks is actually compensation for unobserved comparative advantages of individuals that have endogenously sorted into the different sectors of the economy.

6 Concluding Remarks

Using data from the Survey of Income and Program Participation (SIPP) one finds that the level of risk and the mean level of (log) earnings are positively related . Irrespective of whether risk is of a transitory or permanent nature, the positive association is clear. The question we ask in this paper is whether that estimated correlation is a pure compensating differential or masks other factors, for instance self-selection due to comparative advantage. A standard Roy model places a lot of weight when choosing an industry in a worker's comparative advantage. However, if the latter is driven by unobserved characteristics, it is not possible to tell how much of the observed equilibrium relation between risk and earnings is due to comparative advantage moving into particular industries and how much to a compensating differential (individuals moving away from risky industries). The model we construct assigns a zero compensating differential component to the relationship between the variance of transitory shocks and (log) earnings. On the other hand, the higher earnings observed in industries with higher permanent risk reflect a pure compensating differential.

References

Abowd, J., and M. Stinson (2011): "Estimating Measurement Error in SIPP Annual Job Earnings: A Comparison of Census Bureau Survey and SSA Administrative Data," Working Papers 11-20, Center for Economic Studies, U.S. Census Bureau.

Abowd, J. M., and O. C. Ashenfelter (1981): "Anticipated Unemployment, Temporary Layoffs, and Compensating Wage Differentials," in *Studies in Labor Markets*, NBER Chapters, pp. 141–186. National Bureau of Economic Research, Inc.

Aiyagari, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 109(3), 659–84.

Bewley, T. (1977): "The permanent income hypothesis: A theoretical formulation," *Journal of Economic Theory*, 16(2), 252–292.

Buera, F. J. (2006): "Non-parametric Identification and Testable Implications of the Roy Model," Discussion paper, Northwestern University.

Carroll, C. D., and A. A. Samwick (1997): "The nature of precautionary wealth," *Journal of Monetary Economics*, 40(1), 41–71.

Dillon, E. W. (2012): "Risk and Return Tradeoffs in Lifetime Earnings," Discussion paper, University of Michigan, Department of Economics.

Feinberg, R. M. (1981): "Earnings-Risk as a Compensating Differential," *Southern Economic Journal*, 48(1), 156–163.

Gottschalk, P. T., and M. Huynh (2006): "Are Earnings Inequality and Mobility Overstated? The Impact of Non-Classical Measurement Error," IZA Discussion Papers 2327, Institute for the Study of Labor (IZA).

Guvenen, F. (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1), 58–79.

Heathcote, J., K. Storesletten, and G. L. Violante (2008): "Insurance and opportunities: A welfare analysis of labor market risk," *Journal of Monetary Economics*, 55(3), 501–525.

——— (2009): "Quantitative Macroeconomics with Heterogeneous Households," *Annual Review of Economics*, 1(1), 319–354.

Heckman, J. J., and B. E. Honore (1990): "The Empirical Content of the Roy Model," *Econometrica*, 58(5), 1121–49.

Hoffmann, F. (2010): "An Empirical Model of Life-Cycle Earnings and Mobility Dynamics," Discussion paper, University of British Columbia, Department of Economics.

Hopenhayn, H., and R. Rogerson (1993): "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 101(5), 915–38.

Huggett, M. (1993): "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of Economic Dynamics and Control*, 17(5-6), 953–969.

Kambourov, G., and I. Manovskii (2009): "Occupational Mobility and Wage Inequality," *Review of Economic Studies*, 76(2), 731–759.

Keane, M. P., and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

Kuznets, S., and M. Friedman (1939): "Incomes from Independent Professional Practice, 1929-1936," in *Incomes from Independent Professional Practice, 1929-1936*, NBER Chapters. National Bureau of Economic Research, Inc.

Leigh, J. P. (1983): "Job Choice across Industries when Earnings Are Uncertain," *Quarterly Review of Economics and Business*, 23(3), 54–69.

Low, H., C. Meghir, and L. Pistaferri (2010): "Wage Risk and Employment Risk over the Life Cycle," *American Economic Review*, 100(4), 1432–67.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), 135–146.

Storesletten, K., C. I. Telmer, and A. Yaron (2004a): "Consumption and risk sharing over the life cycle," *Journal of Monetary Economics*, 51(3), 609–633.

Storesletten, K., C. I. Telmer, and A. Yaron (2004b): "Cyclical Dynamics in Idiosyncratic Labor Market Risk," *Journal of Political Economy*, 112(3), 695–717.

A Appendix: Data

We use three Surveys from the Survey of Income and Program Participation (SIPP): 1996, 2001 and 2004. Cleaning of the data until we reach the final sample is identical across the three surveys. We use the public data files from SIPP available at http://www.ceprDATA.org, maintained by the CEPR (Center for Economic and Policy Research) in Washington, D.C. For each Survey, we perform the following steps:

1. Eliminate individuals who simultaneously report missing earnings but positive hours worked.
2. Eliminate individuals who report working in two different industries or those who do not report their industry (self-employed).
3. Eliminate individuals who report being out of the labor force.
4. Eliminate secondary jobs (i.e. we focus on the primary job of the individual).
5. Restrict analysis to individuals older than 22 but younger than 66.
6. Restrict analysis to individuals who are married.
7. Eliminate individuals who do not report complete samples.

We redefine earnings to be unemployment insurance if an individual reports zero hours worked and reports being unemployed. For those individuals who are not unemployed we also eliminated those with very low earnings (less than 600 1996 dollars per month). [To be completed: Report number of observations lost in each step for all three surveys.]

B Appendix: Model Computation

1. The first step is to discretize the distributions for the selection parameters (i.e. the industry-specific skills or productivities). Recall from the model description that we assume normality for the selection parameters: $\theta_j \sim N(\mu_{\theta_j}, \sigma^2_{\theta_j})$. We construct an equi-spaced grid of length N_R for the support of each distribution $G^j_R = \left\{\hat{\theta}^1_j, \ldots, \hat{\theta}^{N_R}_j\right\}$, assuming $\hat{\theta}^1_j = \mu_{\theta_j} - w_R\sigma_{\theta_j}$ and $\hat{\theta}^{N_R}_j = \mu_{\theta_j} + w_R\sigma_{\theta_j}$ and setting $w_R = 4$ and $N_R = 10$.

2. Guess a distribution of masses $\{\mu_j\}_{j=1}^J$ and efficiency levels $\left\{\theta^*_j\right\}_{j=1}^J$ for each of the industries. This yields aggregate employment levels (in efficiency units) for each of the four industries $\{N_j\}_{j=1}^J$. From our technology assumption, the wage rate in each industry is equal to the marginal product of a unit of average efficiency.

3. Given a set of wages $\{w_j\}_{j=1}^J$, we compute the individual's life-cycle problem for each industry and for each value of the industry-specific ability. To solve for the value and policy functions we discretize the space of bond holdings. Current, not future, bond-holdings are required to lie on a grid $G_B = \{b_1, \ldots, b_{N_B}\}$, with $N_B = 100$, and we use linear interpolation to approximate future value functions. We discretize the values of the persistent and temporary shocks, ω and η.[18]. We use $N_P = 5$ points to approximate the persistent component and $N_T = 2$ to approximate the i.i.d component.[19] The construction of the grid and the computation of the transition matrix for the persistent component follow the procedure outlined in Kopecky and Suen (2010).

4. The previous step yields a set of N_R expected value functions for each industry j conditional on a given level of ability,

$$\left\{\left\{\mathbb{V}^k_j = \int V_j(x|\theta_j = \hat{\theta}^k_j)d\Psi_j(x)\right\}_{k=1}^{N_R}\right\}_{j=1}^J.$$

   To increase the degree of accuracy when updating the aggregate efficiency units of labor, we construct a second grid of length $N_{\tilde{R}} > N_R$: $G_{\tilde{R}^j} = \left\{\tilde{\theta}^1_j, \ldots, \tilde{\theta}^{N_{\tilde{R}}}_j\right\}$. The two endpoints of the set $G_{\tilde{R}^j}$ equal the two endpoints of G^j_R. For any point in $G_{\tilde{R}^j}$ not in G^j_R, we linearly interpolate the value functions of its two nearest neighbors in

[18] For computational reasons, we approximate the random walk with a very persistent process; an autocorrelation of 0.999.

[19] Two grid points for the iid component matches the mean and standard deviation, the only two moments that are relevant if the process follows an iid normal distribution.

G^j_R. Denote the value function given an ability $\tilde{\theta}^k_j$ by $\tilde{\mathbb{V}}^k_j$. Finally, the results shown in the main body of the paper assume $N_{\tilde{R}} = 65$.

5. Completing the previous step yields, four each industry, a set of three vectors: a grid $G_{\tilde{R}^j} = \left\{\tilde{\theta}^1_j, \ldots, \tilde{\theta}^{N_{\tilde{R}}}_j\right\}$, a vector of associated probabilities for each element in $G_{\tilde{R}^j}$, $\left\{\tilde{p}^1_j, \ldots, \tilde{p}^{N_{\tilde{R}}}_j\right\}$,[20] and a vector of associated value functions $\left\{\left\{\tilde{\mathbb{V}}^k_j\right\}_{k=1}^{N_{\tilde{R}}}\right\}_{j=1}^{J}$. Denote by $K^* = (N_{\tilde{R}})^J$ the set of all possible combinations of the J ability parameters. In other words there are K^* possible values for the vector $\left\{\tilde{\theta}^{i_1}_1, \ldots, \tilde{\theta}^{i_J}_J\right\}_{i_1,\ldots,i_J=1}^{N_{\tilde{R}}}$. Let $\mathrm{T} : \{1, \ldots, N_{\tilde{R}}\}^J \to \{1, \ldots, K^*\}$ be a mapping that yields a value in the set $\{1, \ldots, K^*\}$ given a J-tuple $\{i_1, \ldots, i_J\}$, where each element $i_1, \ldots, i_J$ belongs to the set $\{1, \ldots, N_{\tilde{R}}\}$. The number $p_{T(i_1,\ldots,i_J)} = p^{i_1}_1 \times \ldots \times p^{i_J}_J$ is the probability attached to the event an individual draws the vector $\theta^{i_1}_1, \ldots, \theta^{i_J}_J$. There are K^* such probabilities and $\sum_{k=1}^{K^*} p_k = 1$. For each J-tuple $\{i_1,\ ldots, i_J\}$ there is also a set of value functions $\left\{\tilde{\mathbb{V}}^{i_1}_1, \ldots, \tilde{\mathbb{V}}^{i_J}_J\right\}$, and an associated index $j^* = argmax \left\{\tilde{\mathbb{V}}^{i_1}_1, \ldots, \tilde{\mathbb{V}}^{i_J}_J\right\}$ that represents the optimal industry choice for that particular vector of industry-specific skills.

6. Once we have computed the optimal industry j^* for each combination of skill-specific vectors, we are ready to update the guesses for the industry populations and the average efficiencies in each industry. The new mass for industry j, μ_j is computed as:
$$\mu_j = \sum_{k=1}^{K^*} \chi_{\{j^*=j\}} p_k,$$
where χ_A is an indicator variable that takes the value 1 if the event A occurs and zero otherwise. In other words, we sum over all probabilities associated with the event "industry j is the optimal choice". To update the average ability value we proceed analogously by computing,
$$\theta^*_j = \sum_{k=1}^{K^*} \theta_j^{\left\{i_j \in \left\{1,\ldots,N_{\tilde{R}}\right\} : k = T(i_1,\ldots,i_J)\right\}} \chi_{\{j^*=j\}}(p_k)/\mu_j.$$

[20]Since we discretize the state-space for the ability distribution, a given probability is computed as
$$\tilde{p}^i_j = \frac{\phi(\tilde{\theta}^i_j, \mu_{\theta_j}, \sigma^2_{\theta_j})}{\sum_{k=1}^{N_{\tilde{R}}} \phi(\tilde{\theta}^k_j, \mu_{\theta_j}, \sigma^2_{\theta_j})},$$
where $\phi(\theta, \mu, \sigma^2)$ is the density of a normally-distributed random variable with mean μ and variance σ^2 evaluated at θ.

The previous expression computes the average ability re-weighting the probabilities to constrain them to sum to 1 (hence the presence of μ_j dividing each of the probabilities). These two quantities can then be used to update wages, and hence compute new value functions, repeating the above steps until the maximum of the absolute values between the guessed efficiencies and the newly computed efficiencies,and the absolute value of the guessed masses and the newly computed value is less than 10^{-4},

$$\max\left\{\left\{|\mu_j^{(i)} - \mu_j^{(i-1)}|, |\theta^{*(i)}_j - \theta^{*(i-1)}_j|\right\}_{j=1}^{J}\right\} < 10^{-4},$$

where $i-1$ and i are two consecutive iterations.

C Appendix: Hourly Earnings

To show that the results are robust to using earnings per hour (as opposed to total earnings), we perform the same empirical analysis as that of the main body of the text. We use only the 1996 sample as the quality of hours seems to be better than in the other two surveys. Except transforming earnings into per-hour units, the steps are the same as those described above. Tables 13 and 14 show the equivalent to Tables 4 and 5 after having re-estimated equations (17) and (24) with earnings per-hour instead of total earnings on as the variable in the left-hand side.

D Appendix: Sectors and Occupations

Our analysis is based on the division of the US labor market into sectors. One may object to that division and claim that a more reasonable one should focus on occupations. We argue, however, that the distributions of workers across industries and occupations are intimately related. The SIPP data set allows us to classify jobs into 14 grand occupational categories. For each worker in an occupational category we observe in which sector she supplies labor. Table 15 shows, for each occupation, the industries in which workers are employed. Specifically, it shows the quantity and identification number of the sectors that concentrate 50% of the individuals in each occupation. Even considering these very broad occupation definitions it is evident that most occupations are concentrated in only a few industries. As an exception, three occupations see workers spread out into more than three industries. These are i) Executive, Administrative and Managerial which are concentrated in Government, Durable Goods Manufacturing, Finance and Medical Services, ii) Administrative Support including Clerical which are concentrated in Transportation, and also in Government, Finance and Medical Services and; iii) Technicians and Related Support which are concentrated in Recreation and Entertainment, Durable Goods Manufacturing, Hospitals and Non Durable Goods Manufacturing. As argued in the main body of the paper, some of these occupations reflect different bars in an occupational (for instance, someone starts as an administrative assistant, and ends up managing a small group of workers). One could focus on a horizontal division of occupations or alternatively focus on industries. We have chosen the latter path.

E Figures

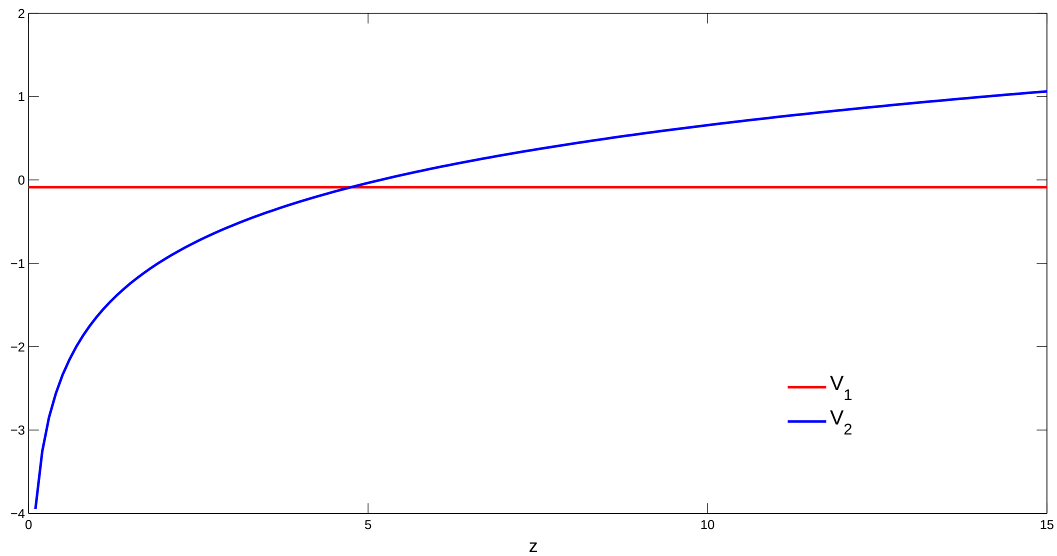

Figure 1: Individual Labor Decision Problem

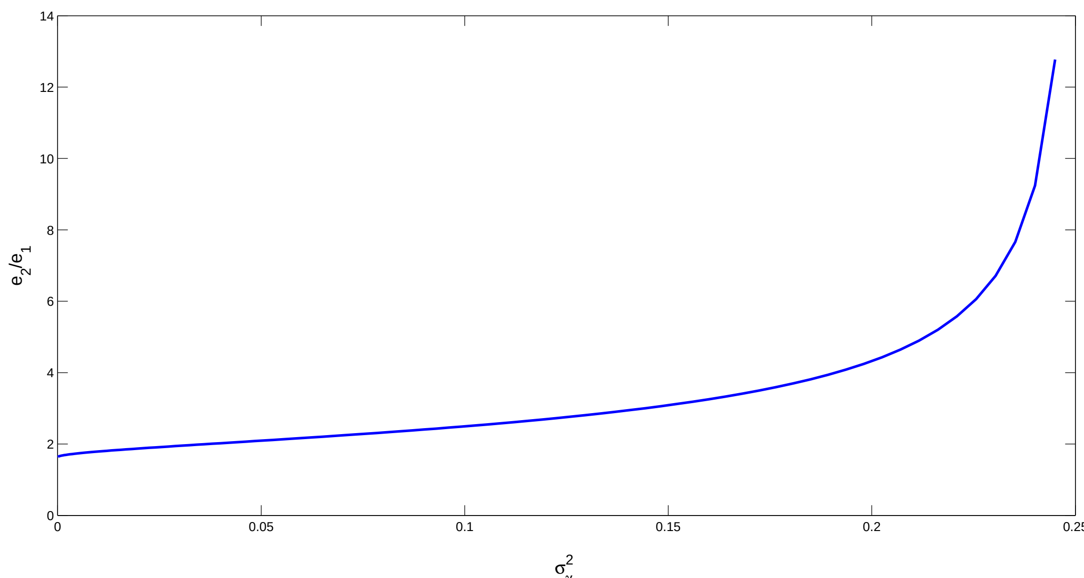

Figure 2: The Price of Risk

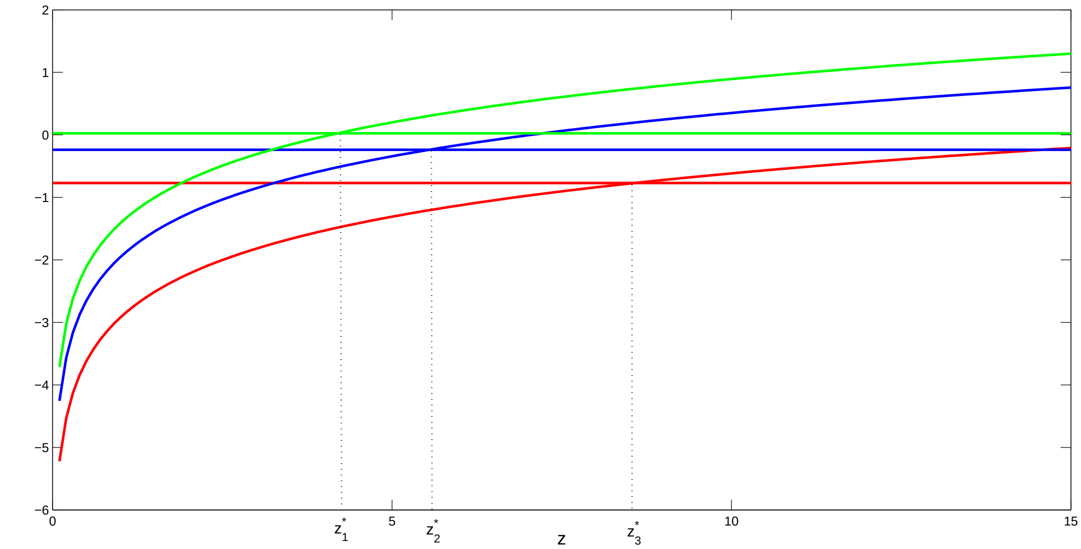

Figure 3: Labor Decision and Variance of Earnings

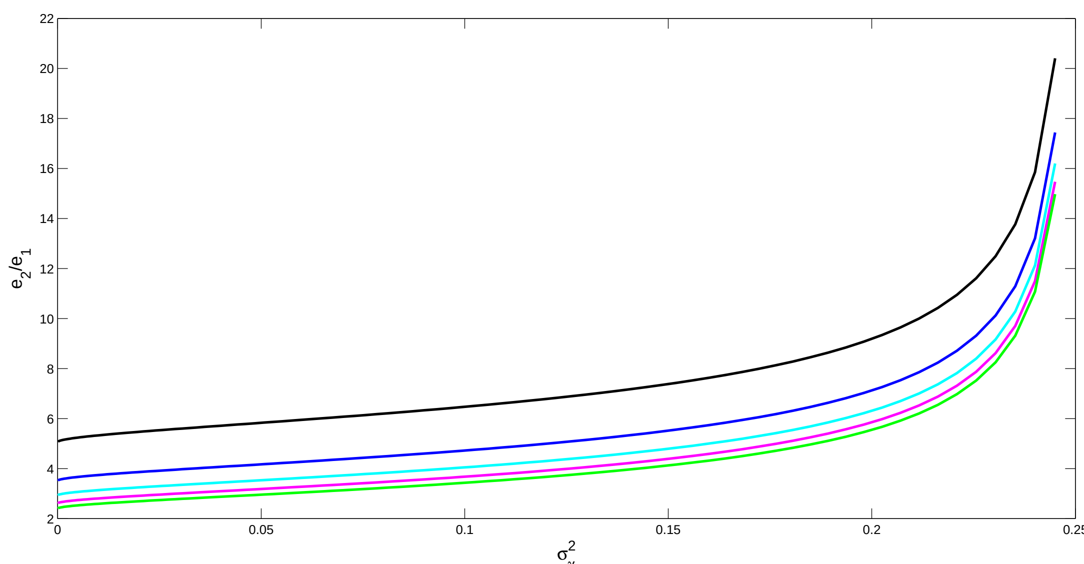

Figure 4: The Price of Risk and Comparative Advantages

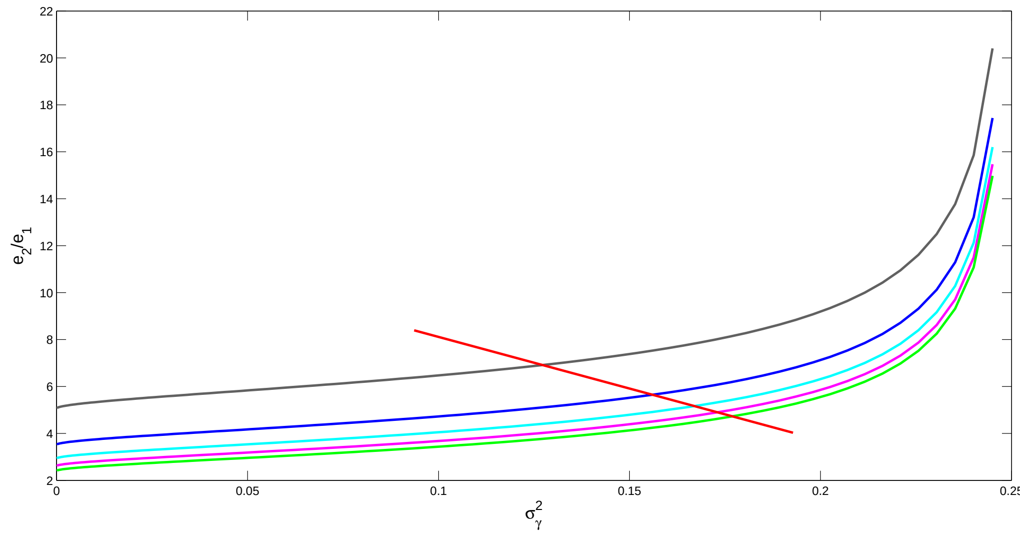

Figure 5: The Effect of Comparative Advantage

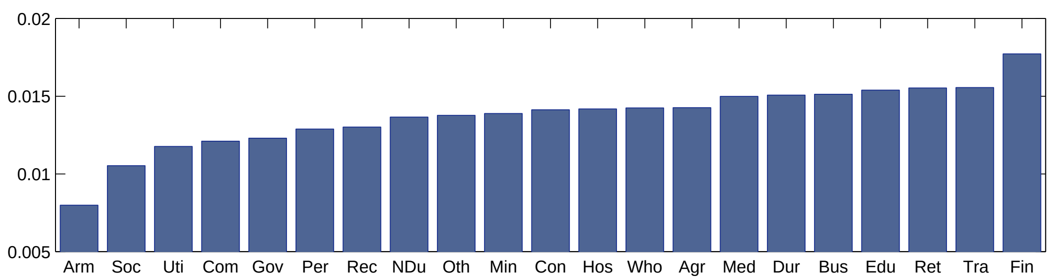

Figure 6: Variance of the Permanent Shock to Earnings

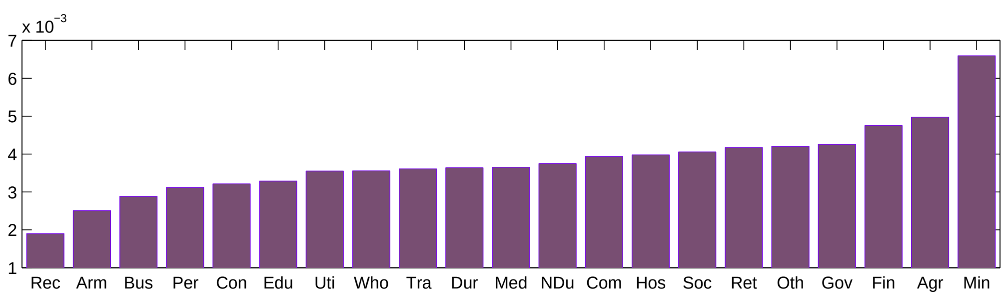

Figure 7: Variance of the Transitory Shock to Earnings

F Tables

Table 1: Variance of Earnings by Industry

	σ_j^2	Ranking
1 Agriculture and Forestry	0.0446	2
	(0.0041)	
2 Mining	0.0545	4
	(0.0063)	
3 Construction	0.0654	9
	(0.0030)	
4 Durable Goods Manufacturing	0.0663	10
	(0.0019)	
5 Non Durable Goods Manufacturing	0.0663	12
	(0.0026)	
6 Transportation	0.0718	18
	(0.0028)	
7 Communication	0.0596	6
	(0.0033)	
8 Utilities	0.0587	5
	(0.0045)	
9 Wholesale Trade	0.0693	14
	(0.0029)	
10 Retail Trade	0.0703	16
	(0.0019)	
11 Finance	0.0901	21
	(0.0039)	
13 Business Services	0.0701	15
	(0.0032)	
14 Personal Services	0.0684	13
	(0.0048)	
15 Recreation and Entertainment	0.0615	8
	(0.0046)	
16 Hospitals	0.0706	17
	(0.0029)	
17 Medical Services	0.0744	20
	(0.0031)	
18 Education	0.0663	11
	(0.0023)	
19 Social Services	0.0531	3
	(0.0028)	
20 Other Services	0.0724	19
	(0.0040)	
21 Government	0.0597	7
	(0.0020)	
22 Armed Forces	0.0395	1
	(0.0070)	

Note: σ_j^2 is the estimate of the variance for the shocks to labor earnings (in logs) for industry j. Bootstrap standard errors are shown in parenthesis. The column Ranking just ranks the industries according to their estimate of the variance.

Table 2: Variance of Earnings by Industry: Transitory and Permanent

	$\sigma^2_{\epsilon,j}$	Ranking	$\sigma^2_{\eta,j}$	Ranking
1 Agriculture and Forestry	0.0143	14	0.0050	20
	(0.0013)		(0.0007)	
2 Mining	0.0139	10	0.0066	21
	(0.0015)		(0.0022)	
3 Construction	0.0141	11	0.0032	5
	(0.0007)		(0.0004)	
4 Durable Goods Manufacturing	0.0151	16	0.0036	10
	(0.0005)		(0.0002)	
5 Non Durable Goods Manufacturing	0.0137	8	0.0037	12
	(0.0005)		(0.0002)	
6 Transportation	0.0156	20	0.0036	9
	(0.0007)		(0.0003)	
7 Communication	0.0121	4	0.0039	13
	(0.0009)		(0.0004)	
8 Utilities	0.0118	3	0.0036	7
	(0.0010)		(0.0004)	
9 Wholesale Trade	0.0142	13	0.0036	8
	(0.0007)		(0.0003)	
10 Retail Trade	0.0155	19	0.0041	16
	(0.0005)		(0.0002)	
11 Finance	0.0177	21	0.0047	19
	(0.0008)		(0.0003)	
13 Business Services	0.0151	17	0.0029	3
	(0.0006)		(0.0003)	
14 Personal Services	0.0129	6	0.0031	4
	(0.0009)		(0.0005)	
15 Recreation and Entertainment	0.0130	7	0.0019	1
	(0.0010)		(0.0004)	
16 Hospitals	0.0142	12	0.0040	14
	(0.0006)		(0.0003)	
17 Medical Services	0.0150	15	0.0037	11
	(0.0007)		(0.0003)	
18 Education	0.0154	18	0.0033	6
	(0.0006)		(0.0002)	
19 Social Services	0.0105	2	0.0041	15
	(0.0007)		(0.0004)	
20 Other Services	0.0138	9	0.0042	17
	(0.0008)		(0.0005)	
21 Government	0.0123	5	0.0043	18
	(0.0004)		(0.0002)	
22 Armed Forces	0.0080	1	0.0025	2
	(0.0013)		(0.0008)	

Note: $\sigma^2_{\epsilon,j}$ and $\sigma^2_{\eta,j}$ are the estimates of the variance of permanent and transitory shocks to labor earnings (in logs), respectively, for industry j. Bootstrap standard errors are shown in parenthesis. The columns called Ranking just rank the industries according to their estimate of the two types of variances.

Table 3: Permanent and Transitory Shocks Across Sectors

		Permanent Shock	
		Below Median	Above Median
Transitory Shock	Below Median	22, 8, 14, 15	3, 9, 4, 13, 18, 6
	Above Median	19, 7, 21, 5, 20, 2	16, 1, 17, 10, 11

This table shows the classification of sectors across four dimensions according to the value of the variance of both the transitory (horizontal) and permanent (vertical) shock below or above the median. Specifically, sectors are classified into the below (above) the median category if its corresponding estimated variance of both the permanent and transitory shocks to earnings is below (above) the median of the estimated variances across sectors. The sectors are represented by their numbers as defined in Table 4

Table 4: Regression Results - Total Risk

Dependent Variable	Earnings Coefficient	Net Earnings Coefficient
constant	-15.31	6.44
	(0.0231)	(0.0000)
female	-0.55	
	(0.0000)	
age	1.11	
	(0.0003)	
age^2	-0.01	
	(0.0000)	
education	0.315	
	(0.0000)	
σ^2	6.06	1.46
	(0.0002)	(0.0104)

The second column shows the estimation results of regressing log earnings by industry the variables listed in its first column. The third column presents the estimation results of regressing the net earnings obtained in a previous step on a constant term (*constant*) and on our estimates for the total variance of the earnings shock(σ^2). For positive (negative) coefficients the values in parenthesis shows the probability that the coefficient is less (bigger) than zero computed by Bootstrap.

Table 5: Regression Results - Permanent and Transitory

Dependent Variable	Earnings Coefficient	Net Earnings Coefficient
constant	-12.21	6.37
	(0.0646)	(0.0000)
female	-0.49	
	(0.0000)	
age	0.95	
	(0.0011)	
age^2	-0.01	
	(0.0011)	
education	0.32	
	(0.0000)	
σ_ϵ^2	9.30	6.87
	(0.18)	(0.0152)
σ_η^2	20.33	16.59
	(0.16)	(0.0771)

The second column shows the estimation results of regressing log earnings by industry the variables listed in its first column. The third column presents the estimation results of regressing the net earnings obtained in a previous step on a constant term (*constant*) and on our estimates for the variance of the permanent shock and transitory shocks to labor earnings (σ_ϵ^2 and σ_η^2, respectively). For positive (negative) coefficients the values in parenthesis shows the probability that the coefficient is less (bigger) than zero computed by Bootstrap.

Table 6: Earnings and Variance of Earnings - 4 Industries

	Mean Earnings	S.D. Mean Earnings	σ^2_ϵ	σ^2_η
Agriculture	6.55	0.3687	0.0141	0.0058
Manufacturing	6.54	0.3869	0.0143	0.0035
Services	6.53	0.3287	0.0141	0.0036
Public Sector	6.50	0.4095	0.0101	0.0034
Correl. w/ Mean Earnings			0.88	0.69

This table shows the earnings statistics that correspond to the aggregation of the 21 industries into 4 main industries. It contains the mean earnings in logs (second column), the standard deviation of mean earnings, the variance of the permanent shock (third column) and the variance of the transitory shock (fourth column).

Table 7: Parameters - Distribution of Pre-Labor Market Skills

	$\mu_{j,\theta}$	$\sigma^2_{j,\theta}$
Agriculture	3.78	0.75
Manufacturing	5.06	0.65
Services	8.51	0.35
Public Sector	5.75	0.65

This table shows the calibrated values for the mean ($\mu_{j,\theta}$) and variance ($\sigma^2_{j,\theta}$) of the distribution of pre-labor market skills for the 4 industries considered.

Table 8: Share of Workers by Industry

	Model	Data
Agriculture	0.03	0.02
Manufacturing	0.05	0.24
Services	0.73	0.65
Public Sector	0.18	0.10

This table shows the model predictions and their data counterpart of the share of workers in each of the 4 industries considered.

Table 9: Wealth to Income Ratios

Agriculture	3.25
Manufacturing	3.53
Services	3.17
Public Sector	1.03
Total Economy	3.04

This table shows the model predictions for the wealth to income ratio in each of the 4 industries considered as well as for the whole economy.

Table 10: Share of Workers by Industry

	Data	Benchmark	Counterfactual
Agriculture	0.02	0.03	0.16
Manufacturing	0.24	0.05	0.22
Services	0.65	0.73	0.21
Public Sector	0.10	0.18	0.48
Correlation with Data		0.92	-0.24

This table shows the model predictions for the share of workers in the benchmark case and in the counterfactual experiment as well as their data counterpart in each of the 4 industries considered.

Table 11: Ratio of Abilities Pre and Post- Sorting

	Mean
Agriculture	4.95
Manufacturing	3.46
Services	1.15
Public Sector	2.52

This table shows the ratio of the mean abilities before (the calibrated values) and after the sorting of workers take place in the model economy.

Table 12: Model Predictions: Mean and Volatility of Earnings

	Benchmark	Counterfactual
Variable	Coefficient	Coefficient
constant	6.39	6.32
Permanent σ_{ϵ}^2	8.51	15.1
Transitory σ_{η}^2	8.38	0.9

The second column shows the estimation results of regressing log earnings by industry to the variance of permanent and transitory shocks in the benchmark model. The third column presents the estimation results of the same regression but using the mean earnings predicted by the model in the counterfactual experiment described in the main body of the text.

Table 13: Regression Results (Earnings Per Hour) - Total Risk

Dependent Variable	Earnings Coefficient	Fixed Effect Coefficient
constant	-9.36	1.42
	(0.0017)	(0.0000)
female	-0.32	
	(0.0011)	
age	0.57	
	(0.0002)	
age^2	-0.01	
	(0.0004)	
education	0.17	
	(0.0000)	
σ^2	5.42	1.67
	(0.0055)	(0.0663)

The second column shows the estimation results of regressing log earnings per hour by industry the variables listed in its first column. The third column presents the estimation results of regressing the net earnings per hour obtained in a previous step on a constant term (*constant*) and on our estimates for the total variance of the earnings shock(σ^2). For positive (negative) coefficients the values in parenthesis shows the probability that the coefficient is less (bigger) than zero computed by Bootstrap.

Table 14: Regression Results (Earnings Per Hour) - Permanent and Transitory

Dependent Variable	Earnings Coefficient	Net Earnings Coefficient
constant	-8.12	1.3928
	(0.0055)	(0.0000)
female	-0.34	
	(0.0041)	
age	0.49	
	(0.0012)	
age^2	-0.01	
	(0.0020)	
education	0.15	
	(0.0008)	
σ_ϵ^2	6.42	3.09
	(0.0509)	(0.0425)
σ_η^2	17.30	0.26
	(0.1338)	(0.5387)

he second column shows the estimation results of regressing log earnings per hour by industry the variables listed in its first column. The third column presents the estimation results of regressing the net earnings per hour obtained in a previous step on a constant term (*constant*) and on our estimates for the variance of the permanent shock and transitory shocks to labor earnings (σ_ϵ^2 and σ_η^2, respectively). For positive (negative) coefficients the values in parenthesis shows the probability that the coefficient is less (bigger) than zero computed by Bootstrap.

Table 15: Distribution of Sectors

Occupation	# Sectors Conc. 50%	Names
1 Executive, Administrative and Managerial	5	$20, 4, 11, 17$
5 Administrative Support including Clerical	4	$20, 11, 6, 17$
3 Technicians and Related Support	4	$15, 4, 16, 5$
8 Services except household and protective	3	$16, 10, 15$
10 Precision Production, Craft and Repair	3	$4, 3, 5$
13 Handlers, Eq Cleaners, Helpers and Laborers	3	$10, 4, 5$
12 Transportation and Material Moving	2	$6, 9$
2 Professional Specialties	2	$17, 15$
4 Sales	2	$9, 10$
7 Protective Services	1	20
9 Farming, Forestry and Fishing	1	1
11 Machine Operators, Assemblers and Inspectors	1	4
14 Soldiers	1	21

This table shows the distribution of occupations across sectors. Specifically, the number of industries (column 3) that concentrates 50% of the workers in the corresponding occupation (column 2). Column 4 lists the identification number of each industry.