Frictional Wage Dispersion: A Puzzle?*

Andreas Hornstein[†] Per Krusell[‡] Giovanni L. Violante[§]

PRELIMINARY AND INCOMPLETE

Abstract

This paper demonstrates that the standard search and matching models of equilibrium unemployment, once properly calibrated, can generate only a tiny amount of frictional wage dispersion, i.e., wage differentials among ex-ante similar workers induced purely by search frictions. The analysis is centered around a specific measure of wage dispersion— the ratio between the average wage and the lowest (reservation) wage paid. We show that in the textbook search and matching models this statistic (the "mean-min ratio") can be obtained in closed form as a function of observable variables, without any parametric assumption on the wage offer distribution. Looking at various independent data sources suggests that, empirically, frictional wage dispersion is larger by a factor of 20. We discuss the extent to which three extensions of the model (risk aversion, volatile wages during employment, and on the job search) can improve its performance.

*We are grateful for comments from Gadi Barlevy, Bjoern Bruegemann, Zvi Eckstein, Chris Flinn, Giuseppe Moscarini, and Philippe Weil, and from seminar participants at the Cleveland Fed, ESSIM 2004 (Cyprus), NYU Macro Lunch, Philadelphia Fed, Penn State, Princeton, USC, Yale, and York University. Greg Kaplan and Matthias Kredler provided excellent research assistance. The views expressed here are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[†]Federal Reserve Bank of Richmond, and CEPR

[‡]Princeton University, CEPR, and NBER

[§]New York University, and CEPR

1 Introduction

The economic success of individuals is largely determined by their labor market experience. For centuries, economists have been interested in studying the determinants of earnings dispersion among workers. The standard theories of wage differentials in competitive environments are three. Human capital theory suggests that a set of individual characteristics (e.g., individual ability, education, labor market experience, job tenure) are related to wages because they correlate to either innate skills or skills cumulated in schools, and on the job. The theory of compensating differentials argues that wage dispersion arises because wages compensate for non-pecuniary characteristics of jobs and occupations such as fringe benefits, amenities, location, and risk. Models of discrimination assume that certain demographic groups are discriminated against by employers and, as such, they earn less for similar skill levels.

Mincerian wage regressions based on cross-sectional individual data proxy all these factors through a large range of observable variables, but typically they can explain at most 1/3 of the total wage variation. A vast amount of variation is left unexplained. In practice, measurement error is large, and the covariates normally available capture only imperfectly what the theory suggests as determinant of wage differentials. However, even if we could "perfectly measure" what these competitive theories require, we should not expect to explain all wage dispersion.

Theories of frictional labor markets which built on the seminal work of McCall (1970), Mortensen (1970), Lucas and Prescott (1974), Burdett (1978), Pissarides (1985), and Mortensen and Pissarides (1994) predict that wage differentials arise among ex-ante similar workers, looking for jobs in the same labor market (e.g., the market for janitors in Philadelphia) because of informational frictions in the labor market, and luck in the search process. We call this type of wage inequality inherently associated to frictions *frictional wage dispersion*.[1]

The canonical search and matching model provides a natural framework to think about frictional wage dispersion. We begin by asking how much frictional wage dispersion the model can generate and we arrive at a surprisingly general implication. We show that in the three standard models of equilibrium unemployment (sequential search model, islands

[1]Mortensen (2005) calls it *pure* wage dispersion.

model, and random matching model) one can obtain the same analytical expression for a particular measure of frictional wage dispersion, the ratio between the average wage and the lowest (reservation) wage paid to employed workers. We call this measure the mean-min (Mm) ratio. The Mm ratio has the desirable property that it does not depend on the shape of the wage offer distribution, but only on a small set of parameters of the model that can be readily calibrated to match well known features of the U.S. economy.

A plausibly calibrated model displays $Mm = 1.036$, i.e., it only generates 3.6% differentials between the lowest paid worker and the average paid worker. The reason is that in the search model good things come only to those who wait, but the data on unemployment duration show that workers do not wait for long. This search behavior of workers rationalizes only a tiny amount of dispersion in the wage distribution they sample from.

The next natural question is: how big is frictional wage dispersion in actual labor markets? Given the features of our model, ideally, one would like to access individual wage observations for ex-ante similar workers searching in the same labor market. This requirement poses several challenges that we address by exploiting three alternative data sources: the November 2000 Occupational Employment Survey (OES), the 1967-1996 waves of the Panel Study of Income Dynamics (PSID), and the 5% IPUMS sample of the 1990 U.S. Census. Overall, from the empirical work we conclude that the observed Mm ratios are at least larger by a factor of 20 than what the model predicts, i.e. empirically $Mm = 1.70$. This large discrepancy between model and data poses a puzzle.

In an attempt to solve the puzzle, we extend the baseline model in three directions: the first is the introduction of risk aversion; the second extension allows for stochastic wage fluctuations during employment and endogenous separations; the third generalization allows for on the job search. This latter modification is the one that shows most promise but, at least in its simplest form, it still falls short of explaining the data.

The paper is still preliminary, so we have not yet comprehensively compared our results to the literature on structural estimation of search models recently surveyed by Eckstein and van den Berg (2005).

The rest of the paper is organized as follows. Section 2 derives the expression for the Mm ratio in three canonical search models and quantifies the implications of the models. Section 3 contains the empirical analysis. Section 4 makes three attempts at rescuing the canonical model. Section 5 outlines three extensions of the model and evaluates them quantitatively. Section 6 concludes the paper.

2 Frictional wage dispersion in canonical models of equilibrium unemployment

The three canonical models of frictional labor markets are the sequential search model developed by McCall (1970) and Mortensen (1970), the islands model of Lucas and Prescott (1974), and the random matching model proposed by Pissarides (1985), and refined by Mortensen and Pissarides (1994).

In what follows, we show that all three models lead to the same analytical expression for a particular measure of frictional wage dispersion, the mean-min ratio, i.e., the ratio between the average wage and the lowest wage paid in the labor market to an employed worker. Then, we explore the quantitative implications of this class of models for this particular statistic of frictional wage dispersion.

2.1 The basic search model

We begin with the basic sequential search model formulated in continuous time. Consider an economy populated by ex-ante equal, risk neutral, infinitely lived individuals who discount the future at rate r. Unemployed agents receive job offers at the instantaneous rate λ_u. Conditional on receiving the offer, the wage is drawn from a well-behaved distribution function $F(w)$ with upper support $w^{\max}$. Draws are i.i.d. over time, and across agents. If a job offer w is accepted, the worker is paid a wage w until the job is exogenously destroyed. Separations occur at rate σ. While unemployed, the worker receives a utility flow b which includes unemployment benefits, value of leisure, and home production net of search costs. Thus, we have the Bellman equations

$$rW(w) = w - \sigma [W(w) - U] \quad (1)$$

$$rU = b + \lambda_u \int_{w^*}^{w^{\max}} [W(w) - U]\, dF(w), \quad (2)$$

where $rW(w)$ is the flow (per period) value of employment at wage w, and rU is the flow value of unemployment. In writing the latter, we have used the fact that the optimal search strategy of the worker is a reservation wage strategy: the unemployed worker accepts all job offers above $w^* = rU$, and has a capital gain $[W(w) - U]$. By simple substitutions, it is immediate to derive the reservation wage equation

$$w^* = b + \frac{\lambda_u}{r+\sigma} \int_{w^*}^{w^{\max}} [w - w^*] \, dF(w).$$

Without loss of generality, let $b = \rho \bar{w}$, where $\bar{w} = E[w | w \geq w^*]$. Then,

$$\begin{aligned} w^* &= \rho \bar{w} + \frac{\lambda_u [1 - F(w^*)]}{r+\sigma} \int_{w^*}^{w^{\max}} [w - w^*] \frac{dF(w)}{1 - F(w^*)} \\ &= \rho \bar{w} + \frac{\lambda_u^*}{r+\sigma} [\bar{w} - w^*]. \end{aligned} \quad (3)$$

where $\lambda_u^* \equiv \lambda_u [1 - F(w^*)]$ is the job finding rate. We obtain an equation relating the lowest wage paid in the economy (the reservation wage) to the average wage paid in the economy through a set of structural parameters of the model.

If we now define the mean-min wage ratio as $Mm \equiv \bar{w}/w^*$, rearranging terms in (3) we arrive at

$$Mm = \frac{\frac{\lambda_u^*}{r+\sigma} + 1}{\frac{\lambda_u^*}{r+\sigma} + \rho}. \quad (4)$$

The mean-min ratio Mm is a measure of frictional wage dispersion, i.e. wage differentials entirely determined by luck in the random meeting process. It measures the *luck premium* between the average worker and the worst-paid worker in a given labor market.

This measure has one important property: it does not depend *directly* in any way on the shape of wage distribution F. Put differently, one can measure Mm without any information on F. The reason is that all it is relevant to know about F, i.e., its probability mass below w^*, is absorbed in the job finding rate λ_u^* that we can directly measure from labor market flows from unemployment to employment.

The mean-min ratio depends on the four parameters $(r, \sigma, \rho, \lambda_u^*)$. The comparative statics with respect to these parameters is informative. A higher discount rate r makes the worker more impatient, reduces the reservation wage and increases Mm. A higher separation rate σ reduces job durations and decreases the value to wait for a better job opportunity (as it would be less durable): this reduces the reservation wage and increases Mm. A lower value of leisure ρ induces agents to accept worse matches, increasing frictional wage dispersion in the economy. Finally, a lower unemployment hazard rate λ_u^* makes the option value of search less attractive, reduces the reservation wage, and pushes Mm up.[2]

2.2 The search island model

We outline a simple version of the island model as done in Rogerson, Shimer, and Wright (2006). Consider an economy with a continuum of islands. Each island is indexed by its productivity level p, distributed by $F(p)$. On each island, there is a large number of firms operating a linear technology in labor $y = pn$, where n is the number of workers employed. In every period, there is a perfectly competitive spot market for labor on every island. An employed worker is subject to exogenous separations at rate σ. Upon separation, she enters the unemployment pool. Unemployed workers search for employment, and at rate λ_u they run into an island drawn randomly from $F(p)$. Every other feature of the model is as before.

It is immediate to see that one can obtain exactly the same set of equations (1)-(2) for the worker, while for the firm in each island, we can write its flow value of producing

$$rJ(p) = pn - wn.$$

Competition among firms drives profits to zero, thus in equilibrium $w = p$. At this point the mapping between the island model and the search model is complete.[3] The search island model yields the same expression for Mm as in (4).

2.3 The basic random matching model

There are three key differences between the search set up and the matching model. First, there is free entry of vacancies. Second, the flow of contacts m between vacant firms and unemployed workers is governed by an aggregate constant returns to scale matching technology $m(u, v)$. Let the workers' contact rate be $\lambda_u = m/u$, and the firm's contact

[2] Also the mean wage $\bar{w}$ is affected in these comparative statics. However, since it declines less than w^*, the comparative statics of the Mm ratio is driven by w^*.

[3] One can also allow firms to operate a constant returns to scale technology in capital and labor, i.e., $y = pk^{\alpha}n^{1-\alpha}$. If capital is perfectly mobile across islands at the exogenous interest rate r, then firms' optimal choice of capital allows to rewrite the production technology in a linear fashion and the equivalence across the two models goes through.

rate be $\lambda_f = m/v = \lambda_u/\theta$, where θ denotes market tightness v/u. Third, workers and firms are ex-ante equal, but upon meeting they jointly draw a value p, distributed with $F(p)$ with upper support $p^{\max}$, which determines flow output on their potential match. Once p is revealed, they bargain on the match surplus in a Nash fashion and determine the wage $w(p)$. Let β be the bargaining power of the worker, then the Nash rule for the wage establishes that

$$w(p) = \beta p + (1-\beta) rU, \tag{5}$$

where rU is the flow value of unemployment. This equation uses the free entry condition of firms that drives the value of a vacant job to zero.

From the worker's point of view, it is easy to see that equations (1)-(2) hold with a slight modification

$$\begin{aligned} rW(p) &= w(p) - \sigma [W(p) - U] \\ rU &= b + \lambda_u \int_{p^*}^{p^{\max}} [W(p) - U] \, dF(p), \end{aligned}$$

i.e., the value of employment is expressed in terms of the value of the match p drawn; similarly, the optimal search strategy is expressed in terms of a reservation productivity p^*. Rearranging these two expressions, we arrive at an equation for the reservation productivity

$$p^* = b + \frac{\lambda_u \beta}{r+\sigma} \int_{p^*}^{p^{\max}} [p - p^*] \, dF(p), \tag{6}$$

where we have used the fact that $p^* = rU$. Substituting (5) into (6), we obtain

$$\begin{aligned} w^* &= b + \frac{\lambda_u [1 - F(p^*)]}{r+\sigma} \int_{p^*}^{p^{\max}} [w(p) - w^*] \frac{dF(p)}{1 - F(p^*)} \\ &= b + \frac{\lambda_u^*}{r+\sigma} [\bar{w} - w^*]. \end{aligned}$$

And using the expression $b = \rho \bar{w}$ into the last equation, we obtain again the formula in (4) for the mean-min ratio. Finally, note that nothing in this derivation depends on the shape of the matching function.

2.4 Quantitative implications for the mean-min ratio Mm

How much frictional wage dispersion can these models generate, when plausibly calibrated? We set the period to one month.[4] An interest rate of 5% per year implies $r = 0.0041$. Shimer (2005a) reports, for the period 1967-2004, an average monthly separation rate σ (EU flow) of 2.0% and a monthly job finding probability (UE flow) of 39%. These two numbers imply a mean unemployment duration of 2.56 months, and an average unemployment rate of 4.88%.

The OECD (2004) reports that the net replacement rate of a single unemployed worker in the U.S. in 2002 was 56%. The fraction of labor force eligible to collect UI is close to 90% (Blank and Card, 1991) which sets the mean replacement rate to roughly 50%. Of course, unemployment benefit are only one component of b. Others are the value of leisure, the value of home production (both positive), and the search costs (negative). Shimer (2005a), weighting all these factors, sets ρ to 41%. As discussed by Hagedorn and Manovskii (2006), this is likely to be a lower bound, but data constraints make it hard to confirm this conjecture. For example, taxes are likely to increase the relative value of ρ since leisure and home production activities are not taxed. Higher values for b will strengthen our argument. We continue by setting $\rho = 0.4$, and we repeatedly return on this key parameter later.

This choice of parameters implies $Mm = 1.036$. The model can only generate a 3.6% differential between the average wage and the lowest wage paid. This number appears very small. What explain the inability of the search/matching model to generate pure wage dispersion? In the model, workers remain unemployed if the option value of search is high. The latter, in turn, is determined by the dispersion of wage opportunities. The short unemployment durations, as in the U.S. data, reveal that agents in the model do not find worthy to wait because frictional wage inequality is tiny. In other words, the essence of search theory is that "good things come to those who wait". If the wait is short, according to the search model it must be because good things are not expected to happen.

The next obvious question is: how big is frictional wage dispersion in actual labor mar-

[4]The Mm ratio has the desirable property of being invariant to the length of the time interval. A change in the length of the period affects the numerator and denominator of the ratio $\lambda_u^* / (r + \sigma)$ proportionately, leaving the ratio unchanged. The parameter ρ is unaffected by the period length.

kets? If, empirically, it is large, then the canonical model fails spectacularly in matching the data. If the data are in line with the model, one would conclude that the role of luck in shaping economic fortunes of individuals in the labor market is negligible.

3 Measurement of frictional wage dispersion

The aim of our analysis is to quantify the empirical counterpart of the mean-min ratio Mm in the model. Given the features of our model, ideally, one would like to access individual wage observations for ex-ante similar workers searching in the same labor market. This requirement poses three major challenges that we address by exploiting three alternative data sources: the November 2000 Occupational Employment Survey (OES), the 1967-1996 waves of the Panel Study of Income Dynamics (PSID), and the 5% IPUMS sample of the 1990 U.S. Census.

First, what is the right definition of a labor market? The most natural boundaries across labor markets are, arguably, geographical, sectoral, and occupational—possibly, combinations of these three dimensions. The PSID sample is too small to construct detailed labor markets. The OES and the U.S. Census allow us to look at the wage distribution in thousands of separate labor markets in the U.S. economy.

Second, differences in annual earnings may reflect differences in hours worked. To avoid this problem, one should focus on hourly wages. However, it is well known that measurement error in hours worked plagues household surveys, and large measurement error will generate an upward bias in estimates of wage dispersion. The OES is an establishment survey where, arguably, measurement error is much lower. PSID and Census data may suffer from measurement error bias. In particular, estimates of the reservation wage through the lowest wage observation is particularly subject to outliers due to reporting errors. As an alternative, we estimate the reservation wage from the 1st, 5th and 10th percentile of the wage distribution. These percentiles are more stable, even though upward biased, estimators of the reservation wage in the empirical wage distribution. We denote the corresponding mean-min ratios as $Mp1$, $Mp5$, and $Mp10$, respectively.

Third, we would like to eliminate all wage variation due to ex-ante differences across individuals in observable characteristics (e.g., experience, education, race, gender, etc.), as well as all wage variation due to heterogeneity in unobservables (e.g., innate ability, value of leisure, etc.). The OES does not provide any demographic information on workers. In the Census data, we can control for a wide set of observable characteristics. When using the PSID, we can exploit the panel dimension to purge also fixed individual heterogeneity from the wage data.

Overall, none of the three data sets is ideal, but each of them is informative in its own way.

3.1 OES Data

The Occupational Employment Statistics (OES) program collects data on workers employed in approximately 1.2 million establishments in order to produce occupational estimates for wages and employment at detailed levels of geography (metropolitan area) and industry (up to 5 digits).[5]

Each establishment reports the average hourly wage paid within each occupation. To the extent that there are within-establishment differences in wages due to luck or frictions, these data underestimate frictional wage dispersion.

We use three different levels of aggregation: the nation-wide data by occupation (3-digit), the occupation × metropolitan area data, and the occupation (3-digit) × industry (5-digit) data. The publicly available data report mean wage, and the 10th, 25th, 50th, 75th and 90th percentiles. The best possible estimate (clearly downward biased) of the mean-min ratio is the ratio between the average wage and the 10th percentile. We exclude all those cells where hourly wage data are not available, and those where the 90th percentile is top coded—a sign that wages are heavily censored at the top in that cell. For the occupation-level data, these restrictions leave 637 cells. We calculate the median Mm ratio across labor markets. The median is preferable to the mean because we found that the empirical distributions of mean-min ratios are very skewed, and we are interested in the mean-min ratio of the wage distribution in a typical labor market. The median $Mp10$ ratio across our 637 occupations is 1.68.

For the classification of labor markets based on 5-digit industry and occupation, we are left with 6,293 cells and the median $Mp10$ ratio is 1.60. Finally, when we define labor markets by metropolitan area and occupation, we have 106,278 cells and the median $Mp10$

[5]See the Appendix for a detailed description of this data source.

ratio is estimated to be 1.48.

As the definition of labor market becomes more refined, wage dispersion falls for two reasons. First, there is less workers' heterogeneity within a specific occupation in a given industry, or in a given geographical area than at the country level. Second, as we keep disaggregating, the number of establishments sampled within each cell falls. For example, for cells defined by occupation and metropolitan area, we have on average only 11 establishment per cell. With such a low number of observations, the lowest wage paid (or the 10th percentile) could be severely upward biased, and in turn the mean-min ratio downward biased.

3.2 PSID Data

Our initial sample comprises of every head and spouse between 20-60 years old in the 1967-1996 waves of the PSID, except for the SEO sample. We then exclude individuals currently in school, self-employed, or disabled, and those with annual hours below 520 and above 5096 to reduce the role for measurement error in hours, which leaves 80,979 individual/year records in the sample.[6] Next, we exclude all individuals whose earnings are top coded or whose hourly wage is below the federal minimum wage, which excludes around 4,141 individual/years observations. At the end of this selection, we are left with 76,848 observations in our final sample.

For every year in the sample period 1967-1996, we run an OLS regression on the cross-section of log hourly wages where we control for gender, 3 race dummies (white, non-white, Hispanics), 5 education dummies (high-school dropout, high-school degree, some college, college degree and post-graduate degree), a cubic in potential experience (age minus years of education minus five), a dummy for union membership, a dummy for marital status, 6 regional dummies, 25 two-digit occupation dummies and interaction between occupation and experience to capture occupation-specific tenure profiles. The objective is to filter out the variation in hourly wages due to observable individual characteristics which are rewarded in the labor market, without incurring in the risk of overfitting. On average, these year-by-year regressions yield an R^2 between 0.42 and 0.45.

[6]French (2005) uses the PSID Validation Study to assess the size of measurement error in hourly wages. He concludes that it accounts for 24% of the standard deviation of log wages. By trimming the hours distribution below 520 and above 5096, we eliminate many outliers due to reporting errors.

Next, we exploit the panel dimension of the data to identify the individual-specific effects in wages. Let ε_{it} be the residual of the first stage for individual $i = 1, ...I$ in year $t = 1, ..., T$. We limit the sample to those who have at least $M_i \geq 10$ wage observations in the panel (49,010 individual/year observations, i.e. 1,633 per year on average) and estimate $\bar{\varepsilon}_i = \sum_{t=1}^{M_i} \varepsilon_{it}/M_i$ for every individual. The vector $\{\bar{\varepsilon}_i\}_{i=1}^{I}$ captures the variation in fixed unobserved individual factors (e.g., innate ability, preference for leisure) which affect wages. Let $\tilde{w}_{it} = \exp(\varepsilon_{it} - \bar{\varepsilon}_i)$. For each year t, we then calculate our indexes of frictional inequality across workers on $\tilde{w}_{it}$.

Figure 1 reports the results. For comparison with the other data sources, we comment on the values for the last part of the sample (the 1990-1996). The ratio between mean wage and lowest wage is $Mm = 4.47$, but the estimate is clearly very noisy. When the reservation wage is estimated from the 1st, and the 5th percentile of the wage distribution, the noise is much reduced and we obtain, respectively, $Mp1 = 2.73$, and $Mp5 = 2.08$. The coefficient of variation of the regression residuals is 0.50.

Controlling for individual effects drastically reduces the estimate of the mean-min ratios, by over a factor of 2. For the period 1990-1996, we estimate $Mm = 3.11$, $Mp1 = 1.90$, and $Mp5 = 1.46$. The coefficient of variation of the residuals net of individual effects falls to 0.25.[7] One should be cautious in interpreting these results though. Through this demeaning procedure we may be eliminating too much variation, including some of what we want to explain. For example, the quality of long-lived matches can confound the estimate of the fixed individual effect.

For comparison with the OES data, we report that the $Mp10$ on the residuals of the first-stage regression equals 1.77 and the $Mp10$ on the residuals net of individual-specific means is 1.32. The corresponding statistics for the OES data all lie somewhere in between.

3.3 Census Data

Our third data source is the 5% IPUMS (Integrated Public Use Microdata Series) sample of the 1990 United States Census. The original data set contains over 12,500,000 person-level observations. To create our sample, first we exclude every person below 20 and above

[7] In passing, we note that Figure 1 is consistent with the views that residual wage dispersion has risen significantly over the period, and that the rise in prices for unobserved innate characteristics is a key component of this phenomenon.

60 years old, as well as every individual currently in school, self-employed, or disabled, which leaves 4,636,759 individual records in the sample. Next, we exclude all individuals who report zero wage income or zero weeks worked over the year, and individuals whose annual earnings is top coded, i.e. higher than $140,000. Finally, we eliminate individuals who report hours worked below 520 and above 5096 and hourly wages below the federal minimum wage ($3.35 in 1989).[8] At the end of this selection, we are left with 3,923,744 individual records in our final sample.

For the log of hourly wage, we run an OLS regression where we control for gender, race, 5 education dummies and a cubic in potential experience. We weight each observation by its Census sample weight. As for the PSID data, this regression has the objective of absorbing the variation in wages due to observable characteristics which are rewarded in the labor market. The regression explains 31% of the total variation in log hourly wages.

Next, we group the (exponent of the) regression residuals by labor markets. Our first definition of labor market is the individual occupation. The Census allows to distinguish between 487 distinct occupations (OCC). Our second definition is the combination of occupation and place of work (for the main job). Our variable for place of work is constructed as follows. Whenever possible, we use the 329 metropolitan areas (PWMETRO), while for rural areas we use a variable (PWPUMA) which defines a geographical area by following the boundaries of groups of counties, or census-defined "places" which contain up to 200,000 residents. Overall, we end up with 799 different geographical areas. For each cell identified as a labor market, we calculate various estimators of the mean-min ratio and we report the median mean-min ratio across labor markets.

In the top panel of Figure 2, we show one example of the wage distribution for Janitors and Cleaners, excluding Maids and House Cleaners, (code 453) in the Philadelphia metropolitan area (code 616). These are the wage residuals of a within-cell regression that controls for the demographics listed above, restricted to those working full time (35-45 hours per week), full year (48-52 weeks per year) to reduce the role of measurement error. Overall, we have 572 observations.

[8]Through validation with CPS data, Baum Snow and Neal (2006) find that a significant fraction of workers report usually working 8 hours per week on the census long form when they actually usually worked 40 hours per week, i.e. they respond as if the question meant to report "usual hours per day". However, this type of measurement error which plagues the 1980 Census is much less frequent in the 1990 census. Moreover, these respondents are excluded from our selection criteria on annual hours.

Table 1: Dispersion measures for hourly wage from the 1990 Census

	Min. obs. per cell	Number of labor mkts	Ratio of mean wage to min.	1st pct.	5th pct.	10th pct.	CV
Occupation		487	4.54	2.83	2.13	1.83	0.47
Occ./Geog. Area	(N≥50)	13,246	2.94	2.66	2.04	1.76	0.41
	(N≥200)	2,321	3.85	2.88	2.13	1.82	0.44
Occ./Geog. Area	(N≥50)	7,195	2.74	2.49	1.92	1.66	0.35
Full time/Full year	(N≥200)	1,117	3.58	2.68	1.98	1.71	0.37
Occ./Geog. Area	(N≥50)	2,810	2.64	2.46	1.92	1.68	0.40
Experience ≤ 10	(N≥200)	406	3.33	2.57	1.97	1.73	0.44
Occ./Geog. Area	(N≥50)	1,152	2.51	2.37	1.98	1.77	0.45
Unskilled Occ.	(N≥200)	191	2.95	2.57	2.08	1.83	0.49
Occ./Geog. Area	(N≥50)	13,246					
Within cell regression	(N≥200)	2,321	3.33	2.66	2.02	1.75	0.42

As reported in the figure, the ratio between the mean and the first percentile is 2.24. In the bottom panel, we display the distribution of $Mp1$ ratios obtained as for the Philadelphia cell, for Janitors and Cleaners across all geographical areas in the U.S. economy for which we have at least 50 observations (131 areas). There are local labor markets markets displaying more and markets displaying less frictional dispersion, but the bottom line seems to be that, even within a very unskilled occupation, even after selecting the sample to minimize the role of measurement error, pure wage dispersion remains large. The median $Mp1$ ratio is 2.20.

Table 1 above reports the results for the entire Census sample. A quick glance at Table 1 reveals that measured frictional wage dispersion is large.

To get at the measurement error issue, we condition our analysis on full time-full year workers who report weekly hours between 35 and 45 and annual weeks worked between 48 and 52: wage dispersion falls with respect to the full sample, but it remains very high. The decline in the coefficient of variation is 16%. For comparison, Bound and

Krueger (1991) compare matched Current Population Survey data to administrative Social Security payroll tax records and find that the measurement error in the standard deviation of log earnings accounts for 7% of the total standard deviation.[9]

To eliminate the importance of individual-specific differences in cumulated skills that we do not capture through experience in the first-stage regression, we condition on workers with less than 10 years of experience, and on a set of very low-skilled occupations, where occupation-specific skills are arguably not very important.[10] Once again, the findings are barely affected.

Finally, we also run the first-stage regressions within each cell, to account for the fact that the role of demographic characteristics in wage determination may be different across occupations. The results of Table 1 remain very robust.

3.4 Summary

To sum up, three independent data sources offered a consistent view of the size of wage dispersion within narrowly defined labor markets. It is large and, in particular, much larger than what the textbook models of section 2 suggest. A review of our estimates yields $Mp5 = 1.46$ from PSID. The PSID estimate could be upward biased because of measurement error, but at the same time the individual wage demeaning could filter out too much variation such as "persistent luck" components. From the Census sample restricted to full-time, full-year workers (where measurement error in hours should be negligible) we have estimated $Mp5 = 1.98$. Given the OES estimate of the $Mp10$, and the fact that the other two datasets suggest that $Mp5$ are roughly 10%-15% larger, we conjecture that the $Mp5$ in the OES data may be around 1.67. An average across the three datasets yields 1.70. This estimate of frictional wage dispersion–based not on the minimum wage observed, but on the 5th percentile–is *19 times larger* than what implied by the model.

In what follows, we use 1.70 as the target for the luck premium between the average and the unluckiest worker. Since we will also need a target estimate of the coefficient of

[9] The standard deviation of the logs has the same scale of the coefficient of variation.

[10] This list includes, inter alia, Launderers and ironers, Crossing guards, Waiters and waitresses , Food counter, fountain and related occupations, Janitors and cleaners, Elevator operators, Pest control occupations, and Baggage porters and bellhops.

variation, we use $cv = 0.30$, an average between the PSID estimate and the estimate on Census sample of full-time, full-year workers.

4 Three attempts to rescue the baseline model

The standard search/matching model of equilibrium unemployment seems to be strikingly unable to match the amount of frictional wage dispersion in the data. In what follows, we try to rescue the model in three different ways.

4.1 Unemployment vs wage dispersion

In defense of the model, one may argue that it is designed to explain unemployment, not wage dispersion. This argument is flawed: in the search model, there is a tight link between the existence of unemployment and the existence of wage dispersion. Unemployment exists because of the option value of searching for better wage opportunities. Let's reverse our logic and suppose that, given the amount of frictional wage dispersion in the data, we want to predict unemployment duration, i.e., use equation (4) and the empirical value of $Mm = 1.7$ to compute the implied value for λ_u^*. We would obtain $\lambda_u^* = 0.011$. In other words, a search model consistent with the amount of wage dispersion in the data predicts an expected unemployment duration of 91 months.

4.2 Implications for other dispersion measures

Admittedly, the mean-min ratio is not a common index of dispersion. One may argue that even though the model fares poorly in terms of this statistic, its performance along more common measures of dispersion, such as the coefficient of variation (cv), could be satisfactory. To answer this question, we need to make further assumptions about the wage offer distribution. Given a parametric specification for this distribution, we can map predicted mean-min ratios into cv's, i.e., we can determine the value of the cv corresponding to a certain value for the mean-min ratio.

The Gamma distribution is a desirable choice because it is a flexible parametric family, and has certain properties that are useful in our application. Let wages w be distributed according to the density

$$g(w; w^*, \alpha, \gamma) = \frac{\left(\frac{w-w^*}{\alpha}\right)^{\gamma-1} \exp\left(-\frac{w-w^*}{\alpha}\right)}{\alpha \Gamma(\gamma)}, \tag{7}$$

with $\gamma, \alpha > 0$, and with $\Gamma(\gamma)$ denoting the Gamma function. The value for w^* is the location parameter and determines the lowest wage observation, α is the scale parameter determining how spread out the density is on its domain, whereas γ is the parameter that determines the shape of the function (e.g., exponentially declining, bell-shaped, etc.).[11]

The mean and standard deviation of a random variable distributed with $g(w; w^*, \alpha, \gamma)$ are given by, respectively, $\mu(w) = w^* + \alpha\gamma$ and $sd(w) = \alpha\sqrt{\gamma}$. Recalling that $Mm(w) = \mu(w)/w^*$, it is easy to obtain a simple relationship between the coefficient of variation $cv(w)$ and the mean-min ratio which only depends on the shape parameter γ,

$$cv(w) = \frac{1}{\sqrt{\gamma}} \left[\frac{Mm(w) - 1}{Mm(w)}\right]. \tag{8}$$

The empirical analysis of section (3) suggest that $cv(w) = 0.30$ and $Mm(w) = 1.7$ are reasonably conservative estimates for the coefficient of variation and the mean-min ratio of the wage distribution within labor markets. From equation (8), this implies $\gamma = 1.88$.

A search model generating a mean-min ratio of 1.036, under the Gamma distribution assumption, would generate a coefficient of variation for hourly wages of 0.025, i.e., 1/12th of the coefficient of variation in the wage data. We conclude that the failure of the model generalizes to more common measures of dispersion as well.

4.3 Alternative parameterizations

To calibrate the pair (λ_u^*, σ), we used the UE flow and the EU flow data. One could argue that λ_u^* represents the job finding rate and σ the separation rate, i.e., they should include also flows from workers out of the labor force into employment and from employment to out of the labor force. Shimer (2005a) reports the monthly separation rate to be 3.5%, and the monthly job finding rate to be 61%. For the same values of r and ρ used in the baseline calibration, we obtain $Mm = 1.038$.

With respect to the interest rate r, we have used a standard value, but it is possible that unemployed workers, especially the long-term unemployed, face a higher effective interest

[11] The Gamma family is very flexible: it includes the Weibull (hence, the exponential) distribution for $\gamma = 1$ and the lognormal, in the limit as $\gamma \to \infty$.

rate if they wanted to borrow. Much less is known about ρ. To assess the robustness of our conclusions to the choice of values for these two parameters, in Figure 3, we plot the pairs (r, ρ) which are consistent with a Mm ratio of 1.7, together with the region of "reasonable pairs" based on our prior. The results are striking and suggest the baseline model cannot be rescued: even for annual interest rates around 40% per year, one would need agents to value one month of leisure the equivalent of *minus* three times the average monthly wage. Positive net values of leisure are consistent with the observed luck premium only for interest rates beyond $1,350\%$ per year.

5 Extensions of the baseline model

We now look at three extensions of the baseline model: risk aversion, stochastic wages during employment, and on the job search.

5.1 Risk aversion

Qualitatively, risk aversion works in the direction of solving the puzzle. Risk averse workers dislike particularly states with low consumption, like unemployment. Compared to risk-neutral workers, ceteris paribus, they lower their reservation wage in order to exit unemployment rapidly, thus allowing Mm to increase.

Let $u(c)$ be the utility of consumption, with $u' > 0$, and $u'' < 0$. To make progress analytically, we assume workers have no access to storage, i.e., $c = w$ when employed, and $c = b$ when unemployed. It is clear that this model will give an *upper bound* for the role of risk-aversion: with any access to storage, self-insurance or borrowing, agents can better smooth consumption, thus becoming effectively less risk averse.

It is easy to see that equation (3) becomes

$$u(w^*) = u(\rho\bar{w}) + \frac{\lambda_u^*}{r+\sigma}\left[E(u(w)) - u(w^*)\right].$$

A second-order Taylor expansion of $u(w)$ around the conditional mean $\bar{w}$ yields

$$u(w) \simeq u(\bar{w}) + u'(\bar{w})(w-\bar{w}) + \frac{1}{2}u''(\bar{w})(w-\bar{w})^2. \tag{9}$$

Take the conditional expectation of both sides of the above equation and arrive at

$$E(u(w)) \simeq u(\bar{w}) + \frac{1}{2}u''(\bar{w})\,var(w). \tag{10}$$

Let $u(w)$ belong to the CRRA family, with θ representing the coefficient of relative risk aversion. Then, using (10) into (9), and rearranging, we obtain

$$Mm \equiv \frac{\bar{w}}{w^*} \simeq \left[\frac{\frac{\lambda_u^*}{r+\sigma}\left(1 + \frac{1}{2}(\theta - 1)\theta cv^2(w)\right) + \rho^{1-\theta}}{\frac{\lambda_u^*}{r+\sigma} + 1} \right]^{\frac{1}{\theta-1}}. \tag{11}$$

It is immediate to see that, for $\theta = 0$, the risk-neutrality case, the expression above equals that in equation (4).

To assess the quantitative role of risk aversion, we use the same parameterization of the risk neutral case, and based on the evidence provided in section 3, we set $cv(w) = 0.30$. Figure 4 plots the pairs of (ρ, θ) consistent with $Mm = 1.70$. For $\theta = 8$, the model can match the data. For $\rho = 0.2$, the Mm in the model matches the data with $\theta = 3.2$. Remember though that we assumed no storage possibility. Plausibly calibrated models of risk averse individuals who have access to a risk-free bond for self-insurance are much closer to full insurance than to autarky (see Aiyagari, 1994).

5.2 Wage shocks during employment

We now extend the basic search model by allowing wages to fluctuate stochastically along the employment spell. Unemployed workers draw wage offers from the distribution $F(w)$ at rate λ_u, but now these wage offers are not permanent. At rate δ, the wage changes, and the worker draws again from $F(w)$. Draws are i.i.d. over time. Separations are now endogenous and will occur at rate $\sigma \equiv \delta F(w^*)$, where w^* is the reservation wage.

The reason why this generalization can potentially generate a larger Mm ratio is that if the wage is very volatile, then the particular value drawn from $F(w)$ by an unemployed worker is not a good predictor of the continuation value of employment. Unemployed workers will therefore be more willing to accept initially low wage offers, which reduces w^* and increases dispersion.

The Bellman equations for employment and unemployment are, respectively,

$$\begin{aligned} rW(w) &= w + \delta \int_{w^*}^{w^{\max}} [W(z) - W(w)]\, dF(z) - \delta F(w^*)[W(w) - U] \\ rU &= b + \lambda_u \int_{w^*}^{w^{\max}} [W(z) - U]\, dF(z). \end{aligned}$$

With respect to equation (1), the value of employment is modified in two ways. First, the endogenous separation rate is now $\delta F(w^*)$. Second, there is a surplus value from accepting a job at wage w which is given by the second term on the right hand side. Exploiting the fact that $rW(w^*) = rU$, integrating by parts and using the fact that $W'(w) = 1/(r+\delta)$, we arrive at

$$w^* = b + \frac{\lambda_u - \delta}{r+\delta} \int_{w^*}^{w^{\max}} [1 - F(z)]\, dz.$$

Now, using the definition of the conditional mean wage $\bar{w}$ and rearranging, we obtain

$$w^* = b + \frac{(\lambda_u - \delta)\,[1 - F(w^*)]}{r+\delta} (\bar{w} - w^*)\,.$$

Therefore, imposing $b = \rho\bar{w}$, and rearranging, we can write the Mm ratio in this model as

$$Mm = \frac{\bar{w}}{w^*} = \frac{\frac{\lambda_u^* - \delta + \sigma}{r+\delta} + 1}{\frac{\lambda_u^* - \delta + \sigma}{r+\delta} + \rho}. \tag{12}$$

As $\delta \to 0$, the Mm ratio converges to equation (4) with $\sigma = 0$, since without any shock during employment every job lasts forever. As $\delta \to \lambda_u$, the wage on the job is i.i.d. and the unemployed worker accepts every offer above b since being on the job has an option value equal to being unemployed.

The parameter δ maps into the degree of persistence of the wage process during employment. In particular, in a discrete time model where $\delta \in (0,1)$, $(1-\delta)$ measures the autocorrelation coefficient of the wage process.[12] Panel data on wages suggest that wages are very persistent, close to a random walk, so plausible values of δ are close to 0.

We can repeat the exercise done in section 4 on the (r, ρ) pair. Given values for (λ_u^*, σ, r) one can search for the values of (δ, ρ) that generate Mm ratios of the observed magnitude. Figure 5 reports the results. Once again, the model is very far from the data, for reasonable values of ρ and of the degree of wage persistence of wage shocks. For example, for $\delta = 0.1$ (annual autocorrelation coefficient of 0.9) the model requires $\rho = -13$. Only for virtually i.i.d. wage shocks ($\delta \simeq 1$) the model would succeed.

5.3 On the job search

We now turn to analyzing on the job search. This extension of the baseline model goes, qualitatively, in the right direction for reasons similar to the model with stochastic wages.

[12] It is easily seen that a discrete time version of this model leads exactly to equation (12).

If the arrival rate of offers on the job is high, workers are willing to leave the ranks of unemployment quickly since they do not forego the option value of search entirely. This property breaks the link between duration of unemployment and wage dispersion that dooms the baseline model.

We generalize the model of section 2 following Burdett (1978). A worker employed with wage $\hat{w}$ encounters new job opportunities w at rate λ_w. These opportunities are drawn from the wage offer distribution $F(w)$ and they are accepted if $w > \hat{w}$. Denoting by w^* the reservation wage, we also assume that no firm would offer a wage below w^*, thus $F(w^*) = 0$.

The values of employment and unemployment become:

$$\begin{aligned} rW(w) &= w + \lambda_w \int_{w^*}^{w^{\max}} [W(z) - W(w)]\, dF(z) - \sigma [W(w) - U] \\ rU &= b + \lambda_u \int_{w^*}^{w^{\max}} [W(z) - U]\, dF(z). \end{aligned}$$

The reservation wage equation becomes

$$w^* = b + (\lambda_u - \lambda_w) \int_{w^*}^{w^{\max}} \frac{1 - F(z)}{r + \sigma + \lambda_w [1 - F(z)]} dz. \tag{13}$$

It is easy to see that, in steady state, the cross-sectional wage distribution among employed workers is

$$G(w) = \frac{\sigma F(w)}{\sigma + \lambda_w [1 - F(w)]}.$$

Using this relation between $G(w)$ and $F(w)$ into the reservation wage equation (13), after simple algebra, we arrive at a new expression for the Mm ratio

$$Mm \simeq \frac{\frac{\lambda_u - \lambda_w}{r+\sigma+\lambda_w} + 1}{\frac{\lambda_u - \lambda_w}{r+\sigma+\lambda_w} + \rho} \tag{14}$$

in the model with on the job search.[13]

Note that, if the search technology is the same in both employment states and $\lambda_w = \lambda_u$, the reservation wage will be equal to b, since searching when unemployed gives no

[13]The "approximately equal" sign comes from one step of the derivation where we have set

$$\frac{r + \sigma}{r + \sigma + \lambda_w [1 - F(z)]} \simeq \frac{\sigma}{\sigma + \lambda_w [1 - F(z)]},$$

a valid approximation since r is a second order term compared to σ.

advantage in terms of arrival rate of new job offers. Indeed, for $\lambda_w > \lambda_u$, unemployed workers optimally accept jobs below the flow value of leisure in order to access a better search technology (during employment).

The crucial new parameter of this model is the arrival rate of offers on the job λ_w. To pin down λ_w, note that average job tenure in the model (see Jolivet et al. (2006) for a similar derivation) is given by

$$\tau = \int_{w^*}^{w^{\max}} \frac{dG(w)}{\sigma + \lambda_w [1 - F(w)]} = \frac{\sigma + \lambda_w/2}{\sigma(\sigma + \lambda_w)} \in \left(\frac{1}{2\sigma}, \frac{1}{\sigma}\right).$$

Since we set the monthly separation rate σ to 0.02, the model can only generate average tenures between 25 and 50 months. The BLS (2006, Table 1) reports that *median* job tenure (with current employer) for workers 16 years old and over, from 1983-2004, was 3.64 years, or 43.7 months. The implied average job tenure would be higher, since tenure distribution are notoriously skewed. However, from the monthly separation rate (3.5%) computed by Shimer (2005a), average tenures appear much shorter, 28.6 months.

We argue that this is a lower bound for average tenure in the U.S. labor market and, in what follows, we ask whether the on the job search model can generate large Mm ratios and, at the same, time be consistent with tenure lengths higher than 28.6 months. It should be clear that the model will have a good chance to generate a large luck premium for low job tenures (or high values of λ_w).

Figure 6 illustrates that λ_w must be 5-10 times smaller than λ_u for the model to generate job durations as long as in the U.S. data, according to the BLS estimates. This means that, once again, the model will produce high Mm ratios only for negative values of ρ.

However, the bottom panel shows that the curve becomes rapidly very flat in the region below tenure lengths of 28 months. Thus, if one accepts the view that job durations are very short and that the offers arrival rate on the job is as large, or even larger, than during unemployment, then the basic on the job search model can exhibit frictional wage dispersion comparable to the data.

5.4 Alternative explanations

Non-pecuniary job attributes: In many jobs, wages are only one component of total compensation. In a search model where a job offer is a bundle of a monetary component and a non-pecuniary component, short unemployment duration can coexist with large wage dispersion, as long as non-pecuniary job attributes are *negatively correlated* with wages so that, the dispersion of total job values is indeed small.

This hypothesis, which combines the theory of compensating differentials to search theory, does not show too much promise. First, certain key non-monetary benefits such as health insurance tend to be positively correlated with the wage, e.g., through firm size. Second, illness or injury risks are very occupational specific and our measures of frictional wage dispersion are within occupations. Third, differences in work shifts and part-time penalties are quantitatively small. Kostiuk (1990) shows that genuine compensating differentials between day and night shifts can explain at most 9% of wage gaps. Manning and Petrongolo (2005) calculate that part time penalties for observationally similar workers are around 3%.

Returns to tenure: Estimates of returns to tenure vary widely. Topel (1991) estimates that 10 years of seniority increase log hourly wages by 24.6%. Altonji and Shakotko (1987) report estimates below 7%. Recently, Altonji and Williams (2004) have reassessed the evidence, concluding that returns to tenure over 10 years could be around 11%, most of them occurring in the first five years of tenure.

To understand how the Mm ratio could be affected by unobserved tenure differences between individuals in the same occupation, recall that median tenure is 3.64 years. If we assume that average tenure is 5 years or greater, then most of the returns to seniority would have been realized by then. We conclude that this factor may account, at most, for 10% differences between the average worker (with average tenure) and the lowest paid worker (with no tenure). Pure wage dispersion still remains very high. Moreover, in Table 1, we showed that Mm ratios remain very large even conditioning on a subgroup of very unskilled occupations, where returns to seniority should be close to zero (janitors, waiters, parking lot attendants, elevator operators, etc.).

6 Conclusions

Search theory suggests that similar workers looking for jobs in the same labor market may end up earning different wages: informational frictions amplify the role of luck in the matching process. However, when plausibly calibrated, the canonical model implies that luck plays virtually no role in the labor market. The data tell a radically different story, which poses a puzzle.

Of course, the measurement of frictional wage dispersion presents some serious challenges since the data are far from ideal. So, one way out of the puzzle is to demonstrate that in the data luck premium is tiny.

We made several attempts to save the model. The most promising extension is the one with on-the-job search. If one believes that average job durations are as short as 2 years, then the on-the-job-search model goes a long way towards matching the data. Even though workers flow data are not inconsistent with such short durations, official job tenure data from BLS suggest average tenures are beyond 4 years.

Other attempts seem less promising. Risk aversion can be successful only if one believes that self insurance is unimportant. A decade of quantitative investigations of "Bewley" models speaks against this possibility. Volatility in wages during employment can be successful only if one believes that wage are as volatile during an employment spell as in the cross-section. Reducing the net value of leisure helps substantially only for large and negative values.

These last two considerations establish an interesting nexus between our work and the Andolfatto-Hall-Shimer puzzle on the inability of the matching model to generate large fluctuations in unemployment and vacancies. On the one hand, the model requires, basically, rigid wages and relative (to the average wage) values of leisure ρ near one to produce volatile vacancies and unemployment rates. On the other hand, it requires hyper-responsive wages and negative relative values of leisure to produce high cross-sectional frictional wage dispersion.

7 Appendix

A: Description of the occupational employment survey

The Occupational Employment Statistics (OES) program collects data on workers in approximately 200,000 non-farm establishments to produce employment and wage estimates for 821 occupations classified based on the Standard Occupational Classification (SOC).

Since November 2003, the program samples 200,000 establishments semi-annually. Before then, it sampled 400,000 once a year. The OES survey is designed to produce occupational wage and employment estimates using six panels (3 years) of data. The BLS Employment Cost Index (ECI) is used to adjust survey data from prior panels before combining them with the current panel's data. The full six-panel sample of 1.2 million establishments allows the production of occupational estimates at detailed levels of geography and industry. Estimates based on geographic areas are available at the National, State, and Metropolitan Area levels. Industry classifications correspond to 3, 4, and 5-digit North American Industry Classification System (NAICS) industrial groups.

The OES survey form sent to establishments defines wages as straight-time, gross pay, exclusive of premium pay. Base rate, cost-of-living allowances, guaranteed pay, hazardous-duty pay, incentive pay including commissions and production bonuses, tips, and on-call pay are included. Excluded are back pay, jury duty pay, overtime pay, severance pay, shift differentials, non-production bonuses, employer cost for supplementary benefits, and tuition reimbursements.

The OES survey groups wages in 12 discrete intervals. In November 2004, the lowest interval was "Under $6.75" and the highest was "$70 and over". Mean hourly wage rate for an occupation equals total wages that all workers in the occupation earn in an hour divided by the total employment of the occupation. The same concept applies to more disaggregated levels such as occupations within metropolitan areas or industries.

The mean wage for each interval is estimated based on occupational wage data collected by the BLS Office of Compensation and Working Conditions for the National Compensation Survey (NCS). The p-th percentile wage for an occupation is calculated by uniformly distributing the workers inside each wage interval, ranking the workers from lowest paid to highest paid.

References

[1] Aiyagari, Rao (1994). "Uninsured Idiosyncratic Risk and Aggregate Saving," Quarterly Journal of Economics, vol. 109(3), pages 659-84.

[2] Altonji, Joseph, and R. Shakotko (1987). "Do Wages Rise With Job Seniority?", Review of Economic Studies, Vol. LIV (3), pages 437-460.

[3] Altonji, Joseph, and Nicholas Williams (2004). "Do Wages Rise with Job Seniority? A Reassessment". Forthcoming in the Industrial and Labor Relations Review.

[4] Blank, Rebecca, and David Card (1991). "Recent Trends in Insured and Uninsured Unemployment: Is There an Explanation?," Quarterly Journal of Economics, vol. 106(4), pages 1157-89.

[5] Baum, Snow, and Derek Neal (2006). "Data Problems and the Measurement of Racial and Gender Wage Gaps Over Time", mimeo.

[6] Bound, John, and Alan Krueger (1991). "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?," Journal of Labor Economics, vol. 9(1), pages 1-24.

[7] Burdett, Kenneth (1978). "A Theory of Employee Job Search and Quit Rates," American Economic Review, vol. 68, pages 212-20.

[8] Burdett Kenneth, and Dale T. Mortensen (1998). "Wage Differentials, Employer Size, and Unemployment," International Economic Review, vol. 39(2), pages 257-73.

[9] Bureau of Labor Statistics (2006); "Employee Tenure Summary," http://www.bls.gov.

[10] Eckstein, Zvi, and Gerard Van den Berg (2005). "Empirical Labor Search: A Survey," IZA Discussion paper, 2005.

[11] French (2005). "The Labor Supply Response to Mismeasured (but Predictable) Wage Changes," Forthcoming in the Review of Economics and Statistics.

[12] Hagedorn, Marcus, and Iourii Manovskii (2005). "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," mimeo.

[13] Jolivet, Gregory, Fabien Postel-Vinay and Jean-Marc Robin (2006). The Empirical Content of the Job Search Model: Labor Mobility and Wage Dispersion in Europe and the U.S.. Forthcoming European Economic Review.

[14] Kostiuk, Peter (1990). "Compensating differentials for shift work," Journal of Political Economy, vol. 98, pages 1054 - 1075.

[15] Lucas, Robert E., and Edward Prescott (1974). "Equilibrium search and unemployment," Journal of Economic Theory, vol. 7 (2), pages 188-209.

[16] Manning, Alan, and Barbara Petrongolo (2005). "The Part-Time Pay Penalty," CEP Discussion Paper No. 679.

[17] McCall, John (1970). "Economics of Information and Job Search," Quarterly Journal of Economics, Vol. 84(1), pages 113-126.

[18] Mortensen, Dale T. (1970). "Job Search, the Duration of Unemployment, and the Phillips Curve," American Economic Review, vol. 60(5), pages 847-62.

[19] Mortensen, Dale T. (2005). Wage Dispersion: Why are similar workers paid differently?, MIT Press.

[20] Mortensen, Dale T., and Christopher Pissarides (1994). "Job Creation and Job Destruction in the Theory of Unemployment," Review of Economic Studies, vol. 61(3), pages 397-415.

[21] OECD (2004); Taxes and Benefits in the United States.

[22] Pissarides, Christopher (1985). "Short-run Equilibrium Dynamics of Unemployment Vacancies, and Real Wages," American Economic Review, vol. 75(4), pages 676-90.

[23] Rogerson, Richard, Robert Shimer, and Randall Wright (2006). "Search Theoretic Models of the Labor Market," Journal of Economic Literature, 43 (4), pages 959-988.

[24] Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander (2004). "Integrated Public Use Microdata Series: Version 3.0," Minneapolis, MN: Minnesota Population Center. http://www.ipums.org.

[25] Shimer, Robert (2005a). "Reassessing the Ins and Outs of Unemployment", mimeo.

[26] Shimer, Robert (2005b). "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," American Economic Review, 95(1), pages 25-49.

[27] Topel, Robert (1991)."Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," Journal of Political Economy, vol. 99(1), pages 145-76.

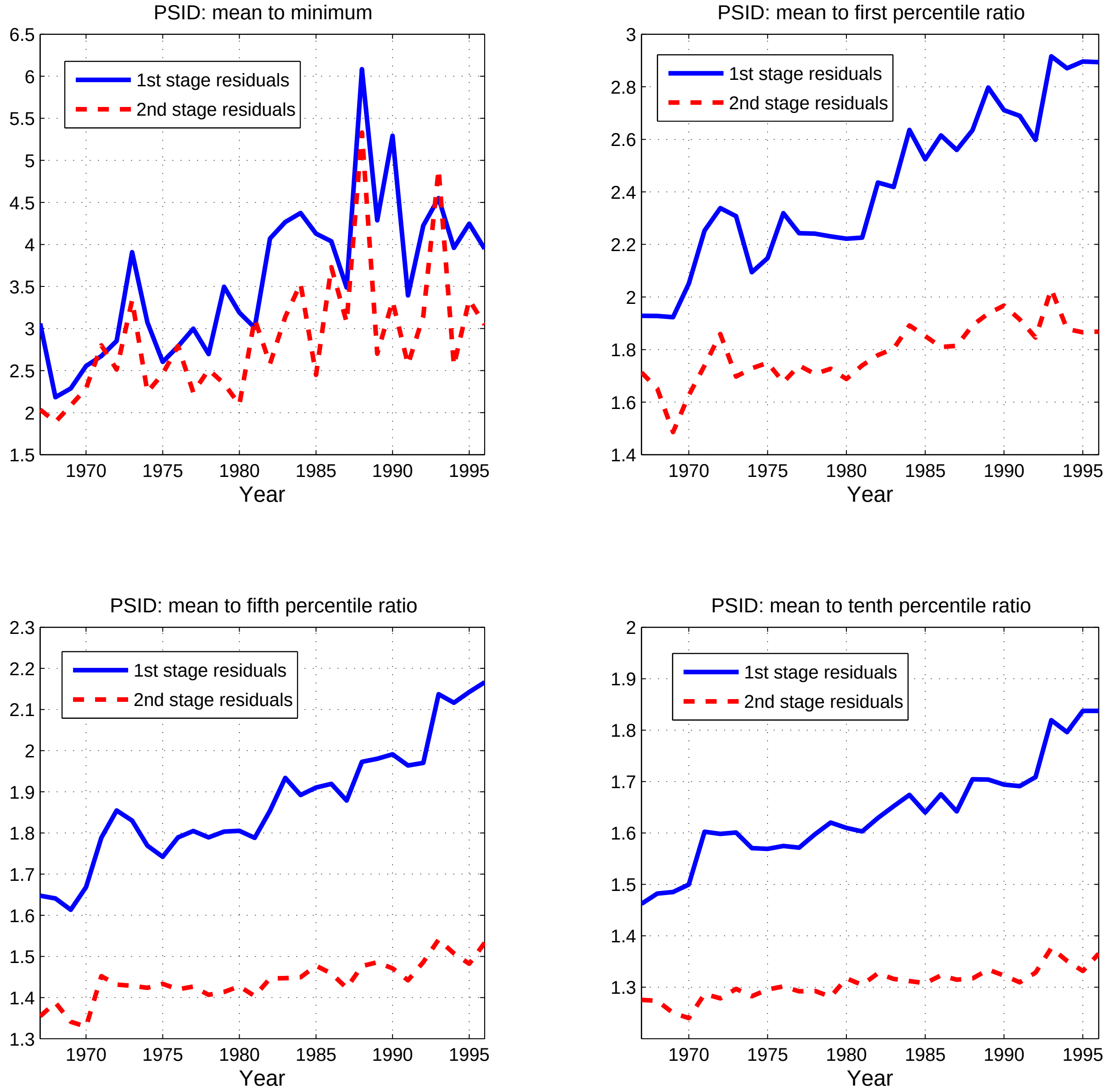

Figure 1: Empirical analysis on PSID data. The first stage residuals refer to the regression on observable covariates. The second stage residuals are the first stage residuals demeaned individually.

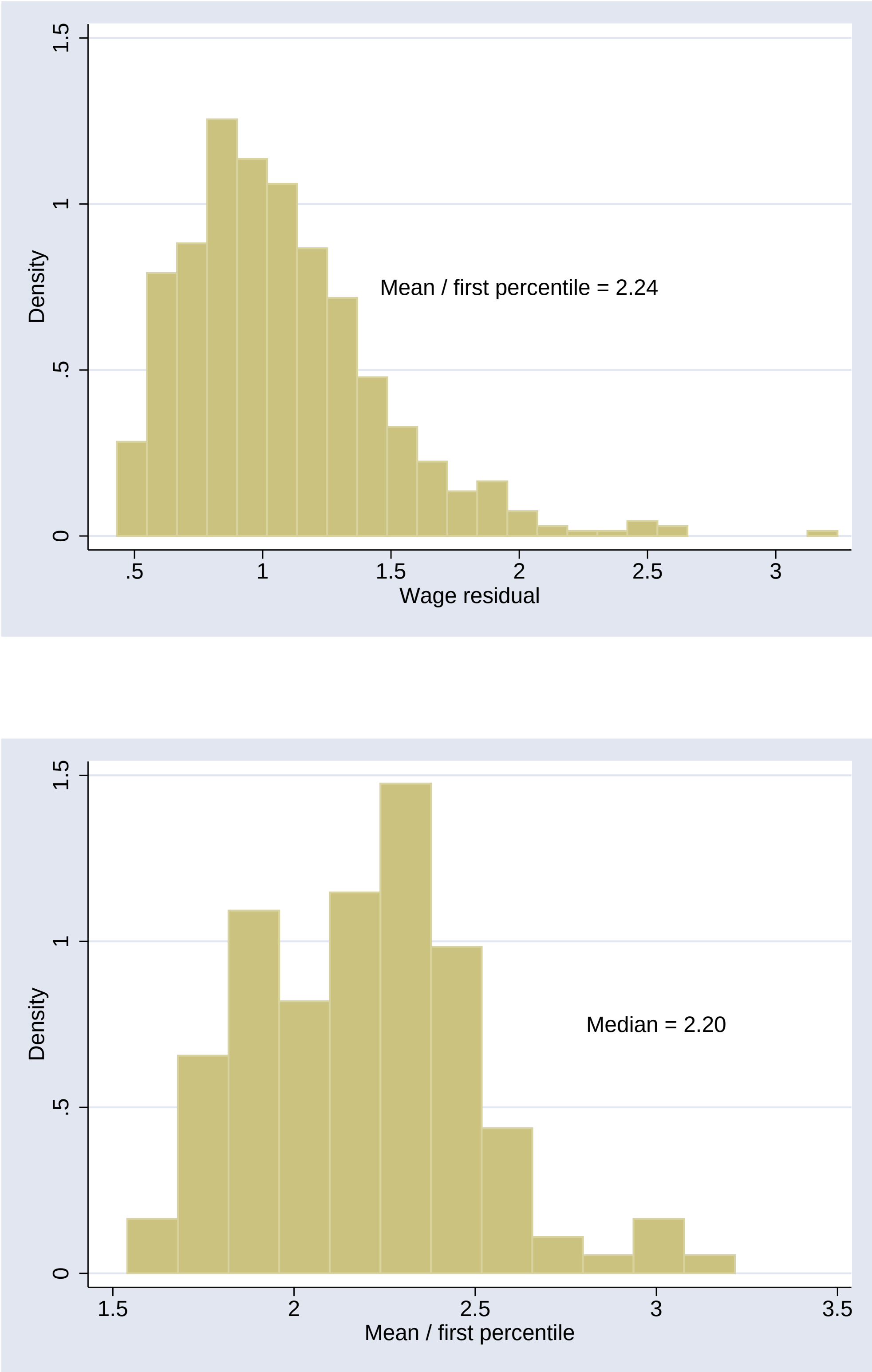

Figure 2: Top panel: Residual wage distribution for full-time, full-year janitors and cleaners in the Philadelphia area. Bottom panel: Distribution of mean-min ratios for full-time, full-year janitors and cleaners across U.S. geographical areas.

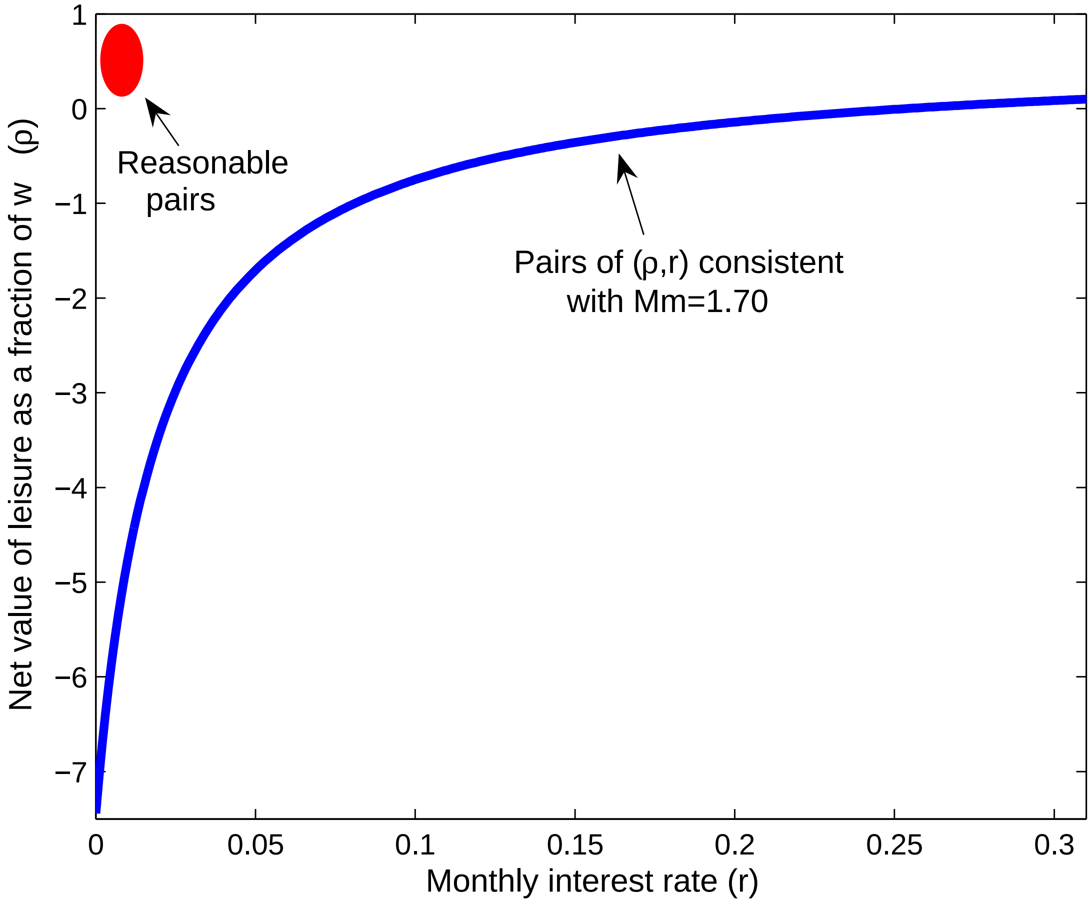

Figure 3: Pairs of the value of leisure and the interest rate that can generate Mm=1.70

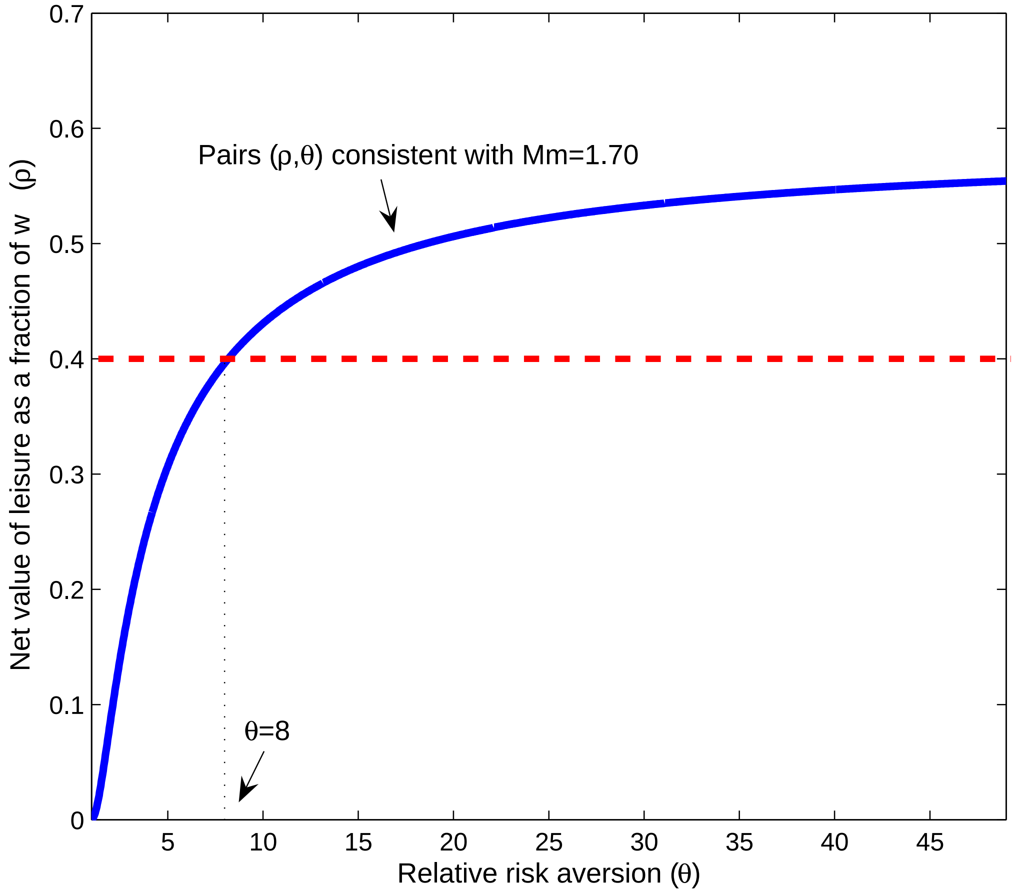

Figure 4: Pairs of the value of leisure and risk aversion that can generate Mm=1.70

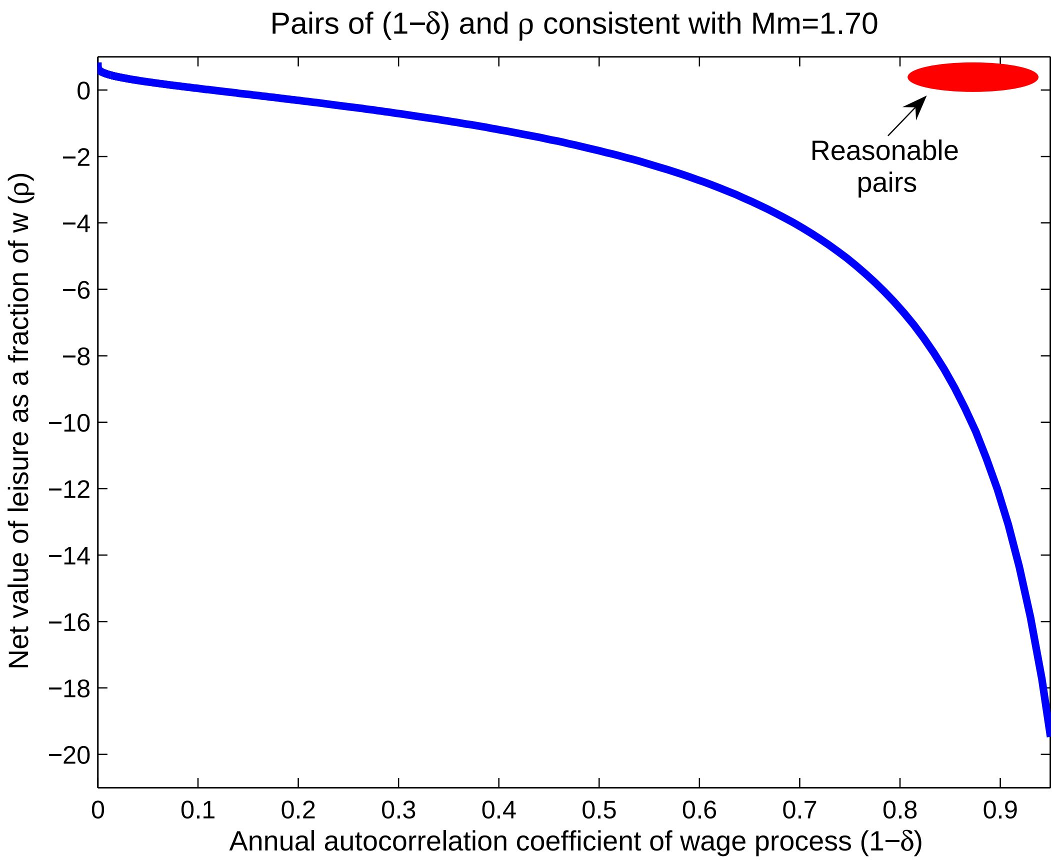

Figure 5: Pairs of the value of leisure and the wage autoregression coefficient that can generate Mm=1.70

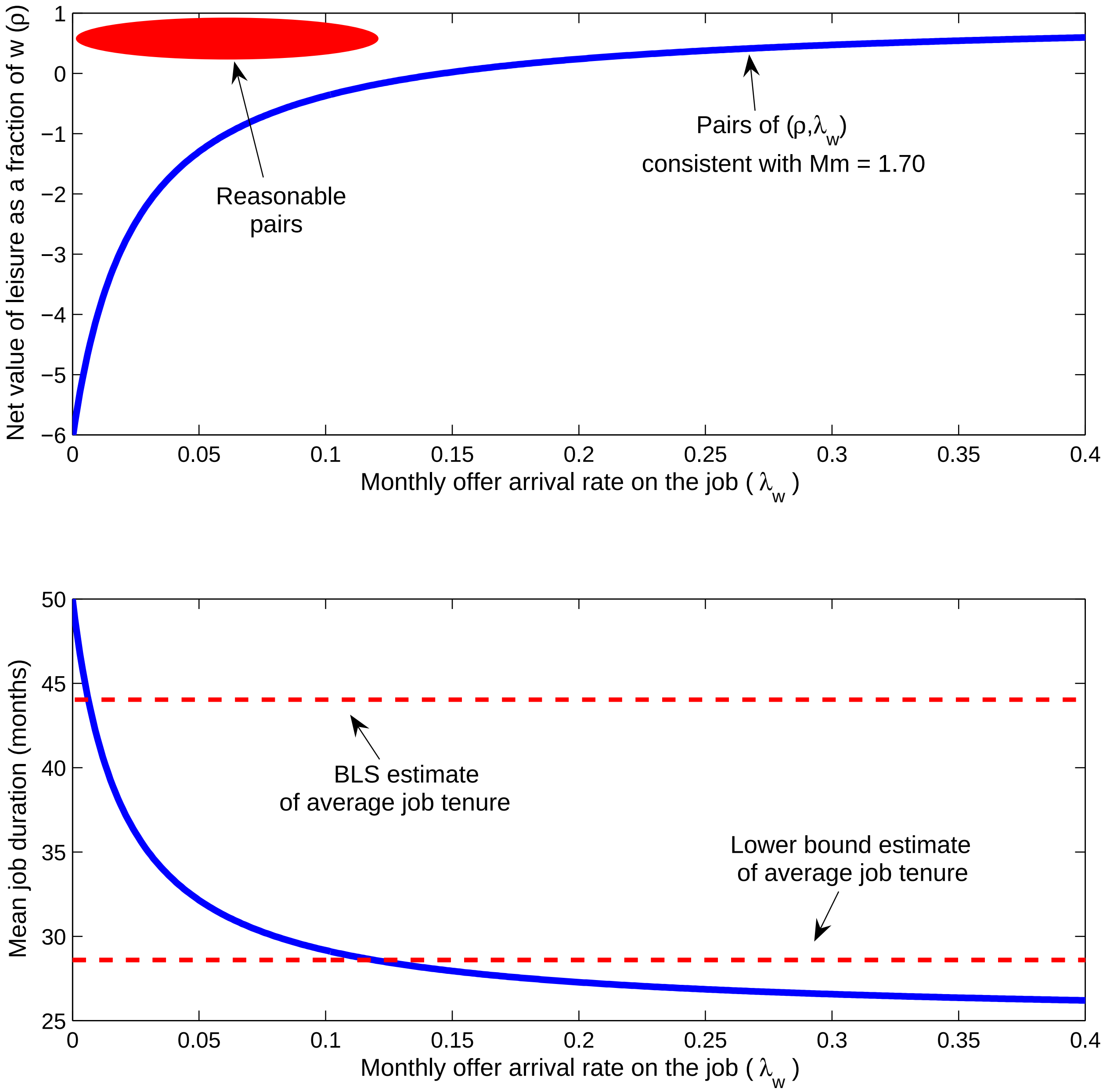

Figure 6: Top panel: pairs of the value of leisure and the job offer rate during employment that can generate Mm=1.70. Bottom panel: mapping between the arrival rate of offers on the job and average job tenure.