Estimating Compensating Wage Differentials with Endogenous Job Mobility*

Kurt Lavetti†
Ian M. Schmutte‡

Preliminary and Incomplete. Please do not circulate.

December 15, 2014

Abstract

We estimate compensating wage differentials for occupational fatality risk using administrative longitudinal matched employer-employee data from Brazil. Our method documents, and corrects for, the presence of bias from endogenous job mobility and non-random assignment of workers to firms that may be correlated with unobserved job characteristics. We find that changes in risk across jobs are correlated with changes in residual wages, so estimates that only control for unobserved worker heterogeneity are biased downward. Controlling for unobserved plant and job-match effects, while allowing for correlation between worker effects, plant effects, and risk, implies compensating differentials that are about 8 times larger than within-worker estimates, and lie between cross-sectional and within-worker estimates. The implied value of a statistical life (VSL) for prime-age male workers, after correction for endogenous mobility, is estimated to be 330,000 reais – equivalent to 42 years employed at the average wage. In addition, our data allow us to measure fatality risk within very detailed industry-occupation cells, alleviating concern about measurement error and aggregation bias that has been highlighted in recent research.

JEL Classifications:
Keywords: Compensating Differentials, Endogenous Mobility

*We are grateful for helpful comments from David Blau, Andrew Friedson, and seminar participants at the Ohio State and the Western Economic Association meetings.

†Ohio State University, Department of Economics. `lavetti.1@osu.edu`
‡University of Georgia, Department of Economics. `schmutte@uga.edu`

1 Introduction

Empirical estimates of compensating wage differentials in labor markets are often of enormous public value. The compensating differential for occupational fatality risk, for example, is frequently used to estimate the value of statistical life (VSL), which affects the cost-benefit analyses of myriad public safety policies, and directly accounts for tens of billions of dollars of federal spending annually in the US. Despite its importance, there are many longstanding estimation problems that are widely acknowledged to have biased nearly every empirical estimate of compensating differentials and labor-market estimates of the VSL to date.

Perhaps the most important of these challenges is the problem of endogeneity bias caused by unobserved worker, firm, and job characteristics that are correlated with both earnings and occupational amenities like fatality rates. Although data have improved greatly since Thaler and Rosen (1976) first wrote about this problem, progress at overcoming these sources of bias has been limited. The main difficulty is that removing bias from unobserved worker characteristics generally requires using variation within-worker, where identification comes from job switches across industries. In within-worker models compensating differentials are estimated based on the ratio of the change in wages to the change in average industry or industry-occupation amenities.[1] However, an as-yet disconnected literature in labor economics, including Abowd et al. (1999), Abowd and Schmutte (2013), Card et al. (2013), and Woodcock (2008), has shown that much of the unobserved variation in earnings across jobs is attributable to changes in firm characteristics and job-match characteristics, which are not portable to new jobs. Moreover, Woodcock (2008) shows that about 90% of the change in earnings associated with job switches is attributable to sorting into higher-paying firms (60 percentage points) and into jobs with higher pure worker-firm match effects (30 percentage points).[2] In within-worker models these earnings changes are wrongly attributed to changes in compensation for occupational risk, causing biased estimates. As a result, as research has progressed from cross-sectional to panel estimation, omitted variable bias from unobserved worker heterogeneity has decreased, but at the likely cost of greatly exacerbating bias due to job search behavior and endogenous assignment of workers to firms.[3]

In this paper we develop a new approach to estimating compensating wage differentials while simultaneously modeling the potentially endogenous assignment of workers to jobs. By accounting for the component of the change in earnings across jobs that is due to changes in unobserved firm and match-specific heterogeneity, we remove bias caused by endogenous

[1]See Brown (1980) and Kniesner et al. (2012).

[2]These estimates are based on US data from the Census Bureau's Longitudinal Employer-Household Dynamics files.

[3]This problem was demonstrated theoretically and through simulations by Hwang et al. (1998), but has not been addressed empirically.

job mobility. We then implement this new strategy to empirically estimate the compensating wage differential for occupational fatality risk, and the associated VSL, using longitudinal matched employer-employee data from the complete census of all formal-sector workers in Brazil between 2003 and 2010. To our knowledge this is the largest database ever used in the estimation of compensating wage differentials.

A second form of improvement that we make relates the measurement of fatality rates. In the US, nearly every study relies on data about fatality rates produced by the Bureau of Labor Statistics, which are released publicly at the two-digit industry level and for restricted access at the two-digit industry by one-digit occupation level. Most international studies use comparable data, also aggregated at the industry and occupation level. Studies that have compared estimates of the VSL based on industry-level aggregate compared to industry-occupation aggregation, including Viscusi (2004), have found measurement error biases on the order of 100%, suggesting that possibility of measurement error within the industry-occupation level could still be biasing estimates substantially. Tsai et al. (2011) estimate the VSL using firm-level risk measures, and finds that variation in risk levels from job switches within industries in Taiwan is larger than the variation from switches across industry. As a result, they find that estimates using industry-level fatality rates are biased downward by an order of magnitude. We build on this progress at alleviating bias due to the aggregation of fatality data by also considering heterogeneity in fatality rates as a function of worker and firm characteristics, such as industry and occupation-specific experience or establishment size. We are able to construct fatality rates at any level of aggregation, including measures that account for worker, establishment, and job heterogeneity, because the data that we use to study earnings also contain a complete census of every occupational fatality in Brazil, linked at the job match level.

Section 2 describes the data and out estimation of fatality rates. Section 3 frames the endogenous mobility problem within the context of identification of hedonic wage models. Section 4 presents an orthogonal match effects model that corrects for unobserved worker, plant and match heterogeneity under the assumption that pure match effects are orthogonal to worker and plant effects, and Section 5 describes estimation results from this model. Section 6 outlines our approach for decomposing the incremental biases that are removed by controlling for unobserved worker, plant, and match effects, each in turn. Section 7 (in progress) presents our correlated random effects model, which relaxes the assumptions of orthogonality between worker and match effects, and between plant and match effects, while allowing correlation between risk and unobserved heterogeneity. Section 8 presents graphical evidence on the distributions of average residuals by deciles of the risk, worker, plant, and match effects distributions, demonstrating the correlations in the residuals caused by endogenous mobility that remain in each of the model specifications. Section 9 concludes.

2 Data and Sample Descriptions

We use matched employer-employee data from Brazil's *Relação Anual de Informações Sociais*, or Annual Social Information Survey (RAIS). The source data play two roles in our analysis: first as a source of information on fatality risk and second to estimate the effect of fatality risk on earnings in the presence of endogenous mobility.

2.1 RAIS Data

RAIS is a census of formal sector jobs. Each year, the Brazilian Ministry of Labor and Employment (MTE) collects data on every formal sector job for the purpose of administering the *Abono Salarial* – a constitutionally mandated annual bonus equivalent to one month's earning. The information in RAIS is provided at the establishment level by a company administrator. In smaller firms and plants, this is likely the owner or plant manager; in larger establishments there may be dedicated personnel who submit the information. Coverage is universal, as employers who fail to complete the survey face mandatory fines and also risk litigation from employees who have not received their *Abono Salarial*.

For every job, the employer reports information on the characteristics of the worker, including a unique identifier that allows us to track the worker from job-to-job. The employer also reports information on the characteristics of the job. For our purposes, the most important job characteristics are the wage, whether the job ended because of a fatal injury on the job, and the worker's occupation. The employer also reports basic characteristics of the plant, including a common identifier and information on plant's industry, location, and the number of employees.

In Brazil a worker is formally employed if he or she has a registered identification number with one of two social security programs: the *Programa de Integração Social* (PIS), or Social Integration Program, or the *Programa de Formação do Patrimônio do Servidor Público* (PASEP), or Civil Servants Equity Formation Program, depending on if the worker is employed in the private sector or the public sector. PIS/PASEP numbers are consistent across workers and follow a worker for life. For firms, formal employment means that the employer contributes to a bank account administered by either *Caixa Econômica Federal*, if registered with PIS, or *Banco do Brasil*, for PASEP workers, covering all worker categories. Formal employers must also have employment contracts for all employees. The most common contract type is the *Consolidação das Leis de Trabalho* (CLT), or Labor Law Consolidation. Other contract types include internships, independent contractors, directorships and government contractors. The Brazilian government defines formal employment with these criteria, and this definition is consistent with definitions used by researchers when studying other Latin American economies (Gasparini and Tornarolli 2009). Formal employment grew steadily in

Brazil during our sample period, from nearly 42 million jobs in 2003 to over 65 million jobs in 2010. Unemployment decreased from 11 percent to five percent, and real wages grew over the period as well. Our sample therefore covers a period of growth and relatively tight labor market conditions.

2.2 Analysis Sample and Variables

The sample we use to estimate the hedonic models is restricted to jobs with positive earnings, hours, and tenure. We also eliminate observations with bad or missing worker and plant identifiers. To eliminate the influence of self-employment on our results, we also restrict attention to jobs in which the plant has at least five workers.

The unit of observation in the raw data is the job-year, where a job is defined by a person-plant combination. For a very small number of such matches, we observe multiple records in the same year. In such cases, we restrict our sample to the observation of a given match with the highest annual earnings. This eliminates about 2 percent of observed match-year records.

For consistency with studies using panel data, in some models we follow Abowd et al. (1999); Woodcock (2008) by restricting our sample to a single job for every worker in every year. We define expected earnings as the product of the average monthly wage rate times the number of months the worker was employed. For each worker, in every year they are employed, we define their dominant job as the job on which they had the highest level of expected earnings. Dropping jobs that are not dominant results in an elimination of approximately 13 percent of all jobs.

The dependent variable is the hourly wage rate. For all jobs, RAIS records the worker's average monthly earnings in nominal Brazilian reais. If the worker is in the job for less than 12 months during the year, the division is adjusted so that the resulting figure represents one month's pay. It is therefore more accurate to think of this variable as measuring a monthly wage rate – a common institutional arrangement in Brazil. For consistency with prior research, we convert to an hourly wage rate. First we calculate a weekly wage rate as the monthly wage rate divided by 4.17. We then calculate the hourly wage rate as the weekly wage rate divided by the contracted weekly hours, which are also reported for every job. We report all wages and earnings in 2003 Brazilian reais. Conveniently, the inflation rate and exchange rates over this period are such that one Brazilian reais in 2003 is approximately equal to one 2010 dollar.[4] While this does not reflect the relative costs of consumption, it provides a point of reference when interpreting magnitudes.

[4]One Brazilian real in 2003 is worth approximately 1.5 Brazilian reais in 2010. Likewise, in 2010, one U.S. dollar was worth 1.66 Brazilian reais.

Human capital controls in the model include indicators for educational attainment and labor market experience. We calculate labor market experience by adding to initial experience the number of months in which a worker had any paid employment over each year. The initial experience is computed as the maximum of tenure in the first observed job or potential experience; whichever is largest. For each job, the data report the date of hire. Hence, even for the first in-sample job, we have an accurate measure of tenure on that job. We also include controls, when appropriate, for gender and race. Because individual characteristics are reported by the employer, they are subject to change as workers move from job-to-job. These apparent discrepancies do not pose a problem in our analysis since these variables just appear as controls. Cornwell, Rivera, and Schmutte (2014) provide evidence that discrepancies in employers' reports of worker characteristics are associated with other unobserved determinants of earnings, so we leave these variables in as reported. We also control for industry, occupation, and plant location (state or municipality). Table 1 presents descriptive statistics of key variables.

2.3 Measures of Fatality Risk

When a job ends during the year, the employer is required to report the cause of separation. The exact cause is important in determining the forms of severance compensation to which the worker is entitled. The employer can choose from 23 different options, including three that cover work-related fatalities (see Appendix Table **??**). We therefore observe the complete population of formal sector jobs, along with the duration of those jobs and the number of hours contracted. We also observe whether each of those jobs ended because of a fatal injury. This puts at our disposal a census of fatal occupational injuries from which we can construct measures of fatality risk along different observable dimensions.

Our analysis is based on measures of the fatality risk for a worker's two-digit industry by three digit occupation group. We distinguish 11,440 industry-occupation groups. We construct the fatality risk within each cell from the raw RAIS micro-data as the number of fatal injuries per 100,000 full-time full year workers. This is equivalent to the approach taken by the Bureau of Labor Statistics in reporting fatal injury rates since 2007 [5] One relative advantage of our data is that we observe both the number of months a job lasted as well as the number of contracted weekly hours. By contrast, the BLS fatality rates are scaled by average hours at work from the CPS.

[5]See `http://www.bls.gov/iif/oshnotice10.htm` for a description of how and why the BLS constructs hours-based fatality rates.

Within a cell, c, we construct the fatality rate, a_c as

$$a_c = \frac{F_c}{(H_c/2,000)} \times (100,000). \tag{1}$$

The numerator, F_c is the number of fatal injuries in cell c. The denominator is the number of full-time full-year equivalent jobs, assuming a 40 hour work week and a 50 week work year. H_c is the total number of contracted hours worked over the year. [6] For each job, j, in the cell c, we count the number of hours worked as $H_i = (MonthsWorked/12) * 50 * (Hours/Week)$. H_c is the sum of H_i over all i. Finally, we inflate the count by 100,000.

Given the level of disaggregation, we construct fatality rates based on a three-year moving average. For example, the fatality rates for 2005 are constructed using fatality counts and hours across all jobs from 2003, 2004, and 2005. The rates for 2006 are computed from data for 2004, 2005, and 2006, and so on. We do so for comparability across studies, and also to smooth out fluctuations in the annual fatality rates. We explicitly follow Kniesner et al. (2012) in assuming our measures of the fatality rate are good proxies for the subjective risk assessments of workers and their employers. It is possible that subjective risk assessments vary systematically from the measured fatality rates. However, our disaggregation by narrow industry and occupation give us much more variation in the fatality risk than has been available in previous studies. We therefore expect measurement error to be less of an issue.

Table 2 reports fatality rates by year in aggregate and across fifteen major industries and nine major occupations. The overall fatality rate is 4.92 fatalities per 100,000 full-time full-year jobs. Over the sample period, the fatality rate declines. In 2003, the fatality rate is 5.91 and hits a maximum in 2004 at 6.313 before declining to 4.202 in 2010. By comparison, the fatality rate in the U.S. was 3.7 per 100,000 full-time full-year jobs. The decline in fatality risk in Brazil appears across all industries and occupations. Fatal injuries are highly concentrated in specific types of job. The fatality rate is highest in the Agriculture, Mining, Construction, and Transportation sectors. Among occupations, the fatality rate is highest among Production I workers, and lowest among Professionals.

3 Identification Problems Caused by Endogenous Mobility

Rosen (1974) describes a fundamental identification problem related to unobservable worker and firm characteristics. To characterize this problem, consider workers with heterogeneous preferences:

$$U(c, z, x, \epsilon)$$

[6] Changes in the definition of full-year work will only affect the scale of our fatality rates. We chose a definition close to the BLS definition, although in Brazil full-year work may be closer to 48 weeks.

over consuption, c, a non-wage job amenity (or disamenity), z, who also have observable and unobservable characteristics, x and ϵ. Similarly, characterize firm accoring to their profit functions:

$$\Gamma(z, y, \eta)$$

where firms have some technology for providing z to workers, and have observable and unobservable characteristics y and η.

In equilibrium, the hedonic envelope function will be the set of tangency points between the indifference curves of workers and offer curves of firms. These tangencies are also affected by assortative matching, potentially based both on observable characteristics like amenities as well as unobservable characteristics of workers and firms. For example, the literature in labor economics that studies matching between workers and firms suggests that there is a slightly positive correlation between unobserved pure worker effects and unobserved pure firm effects based on wage decompositions. These unknown heterogeneity components introduce potential sources of bias into the estimation of hedonic wage models. Under certain assumptions about competition and frictionless matching of workers and firms, the slope of the equilibrium hedonic function will equal a weighted average of the curvatures of firms' isoprofit functions workers' indifference curves. The weights can be expressed in terms of the relative variances of the distributions of ϵ and η, the unobserved heterogeneity parameters.[7] Given this, it is possible to identify the preference and technology functions separately under these assumptions. However, if search is costly the problem becomes substantially more difficult.[8]

Although Rosen's identification strategy is more complicated than most empirical strategies used to estimate compensating wage differentials, these models all implicitly follow from Rosen's work. Both Rosen (1974) and Ekeland et al. (2004) describe several unique cases in which workers' preferences and firms' technology are directly identified. The first is when firms have identical technology, which causes η to drop out of the hedonic price function, and the variation in worker preferences sweeps out the firms' common offer curve. If data are also available on changes over time in profit functions, then the cross-sectional and intertemporal variation can be used to estimate both the isoprofit functions and indifference curves, which characterize all of the primitives of the problem. A second unique case occurs when workers have identical preferences, in which case cross-sectional observations reveal workers' preferences directly.

This second case has been the implicit assumption throughout most of the history of the compensating wage differentials literature. However, labor economics more generally

[7]See Ekeland et al. (2004).

[8]See Hwang et al. (1998).

has shown that the assumptions required by each of these special cases are very unlikely to hold. A large and growing literature, including Abowd et al. (1999), Card et al. (2013), and Woodcock (2008), has documented that good jobs and bad jobs exist, and demonstrated the importance of job search and the assignment of workers to firms based in part on unobserved heterogeneity, in determining wages. Both theory and empirical evidence support the cause for concern that these unobserved components are correlated with non-wage job amenities, and bias compensating wage differential estimates. More directly, Hwang et al. (1998) show through similation that when job search matters, so that firm and match effects differ across jobs, then this basic hedonic wage model will produce biased estimates of the marginal willingness to accept fatal risk.

3.1 Estimation of Hedonic Wage Models

The general form of the hedonic wage model most commonly used in the literature is:

$$w_{ij} = x_{ij}\beta_1 + a_j\beta_2 + \epsilon_{ij} + \nu_i + \mu_j + \eta_{ij} \tag{1}$$

where w_{ij} is the log wage of worker i at firm j, x_{ij} contains observable characteristics of both the worker and firm (or industry), and a_j is the fatality rate (or other non-wage job amenity of interest). The error term consists of ϵ_{ij}, a random disturbance, and ν_i, μ_j, and η_{ij}, which represent unobserved worker, establishment, and match-level heterogeneity, respectively. Viscusi and Aldy (2003) reviewed the literature on the estimation of compensating wage differentials for occupational fatality risk in US labor markets, and of the 32 studies in their review all but one relied upon this basic cross-sectional model for identification, and every study used annual industry-average fatality rates as a measure of a_j.

The intention of researchers estimating Equation (1) is to control for enough observable characteristics of workers in x_{ij} that the model resembles Rosen's second special case. However, the first problem with this approach is that a large body of literature in labor economics, including Abowd et al. (1999), shows that workers' unobserved characteristics explain a substantial share of residual variation in wages. To the extent that workers whose unobserved characteristics cause them to earn higher wages spend some of their additional wages on better job amenities, β_2 in Equation (1) is biased.

Similarly, the large unexplained inter-industry and inter-firm wage differentials found in many studies, such as Card et al. (2013), Krueger and Summers (1988), Abowd et al. (1999), and Abowd et al. (2012), suggest that unobserved firm and industry characteristics explain a substantial amount of the residual variation in wages after conditioning on worker effects.[9]

[9] Krueger and Summers (1988) estimate the standard deviation of industry wage differentials to be greater than 10% after controlling for occupation, human capital, and demographic factors. Groshen (1991) estimates

If it is costly for workers to search for jobs with high values of μ_j and η_{ij}, then job search can bias estimates of workers' preferences. Using fixed effects specifications to control for unobserved residual establishment and job-match heterogeneity is generally not possible in hedonic wage models because key job amenities like fatality rates are typically measured at the industry level, and there is generally no exogenous variation in amenities within the job level with which to trace out workers' preferences. Most researchers have instead focused on finding data sources with more complete observable characterizations of workers and firms, but progress on this front has been insufficient to alleviate concerns about omitted variable bias.[10]

The main alternative approach in the literature, including Brown (1980) and Kniesner et al. (2012), is to use panel data from workers who switch jobs in order to remove bias caused by worker effects, ν_i. Identification in these models comes from changes in wage-risk pairs, which requires limiting the identifying variation to changes in wages caused by job switches across industries. A fixed effects specification that includes worker effects removes the pure worker effect and the employment-duration weighted average of the pure establishment effects for the establishments at which each worker was employed, leaving residual unobserved establishment and match heterogeneity, which are the unobserved components of the heterogeneity that gives rise to gains from job search.

However, variation within-worker exacerbates the component of estimation bias that is caused by endogenous job mobility, relative to cross-sectional models. Woodcock (2008) estimates that the rate of earnings growth among workers who experience job-to-job transitions is about three times larger than that of job stayers, and among switchers about 60% of the differential earnings growth is due to sorting into higher paying firms, while 29% is due to sorting into jobs with larger pure match effects. When the identifying variation is limited to within-worker job switches, the relative importance of sorting on the determination of wages increases substantially. Abowd et al. (2010) devise a test for endogenous mobility, which they define as a systematic relationship between pure match effects and wage rates following subsequent job transitions. Abowd et al. (2010) implement this test and strongly reject the null hypothesis of exogenous job mobility, finding that workers with more negative pure match effects are more likely to switch jobs and job transitions tend to increase wages. Following this definition of endogenous mobility, the specific bias caused by job-switches that is problematic in longitudinal hedonic wage models is that the sum $\mu_j + \eta_{ij}$ is likely to change when a worker switches jobs, and this change is wrongly attributed to differences in compensating wage differentials."

the standard deviation of establishment wage differentials to be about 14%.

[10]See Viscusi and Aldy (2003) and Bonhomme and Jolivet (2009).

4 Orthogonal Match Effects Model

Building on Woodcock (2008), we first estimate a reduced-form two-stage model of compensating wage differentials that accounts for unobserved worker, establishment, and job match heterogeneity. To do this, we estimate Equation (2)

$$w_{ijt} = x_{it}\beta + \gamma a_{ijt} + \Phi_{i,J(i,t)} + \epsilon_{ijt} \tag{2}$$

where w_{ijt} is the log earnings of worker $i = 1, \ldots, N$ who is employed at establishment j at time t, x_{it} is a vector containing time-varying characteristics of worker i including experience and year effects, a_{ijt} is the fatality rate of the job at which worker i is employed at time t (note that we denote the fatality rate with subscripts ijt to reflect the fact that we use measures of the fatality rate that vary both over time and within the firm level, such as industry-occupation averages). Following the wage decomposition model proposed by Abowd et al. (1999) (AKM), $\Phi_{i,J(i,t)}$ is the worker-establishment match effect, which absorbs both the pure person effect and the establishment effect.

This model is similar to Kniesner et al. (2012) in controlling for match effects. The identification of the VSL from γ is based on variation in the fatality rate that occurs within the job-match level, which can only come from changes over time in average fatality rates. A major concern with this identification strategy is the salience of these changes. Whereas job switches are frequently associated with large changes in risk levels, frequently on the scale of several times the mean fatality rate, changes over time tend to be extremely small by comparison. If workers cannot readily observe the change in the fatality rate and use that information to negotiate a change in earnings, then the basic assumptions of Rosen (1974) model do not hold. Instead we prefer to make use of the larger source of variation that comes from job switches, but also account for the fact that a great deal of the change in wages associated with a job switch comes from other differences besides just changes in the fatality rate.

To do this we estimate a second stage model, Equation (3), in which the dependent variable is $R_{ijt} \equiv \widehat{\gamma} a_{ijt} + \widehat{\Phi_{i,J(i,t)}} + \epsilon_{ijt}$. The variable contains all of the conditional variation in the wage that is due to either the fatality rate or to unobserved job match heterogeneity. We also include the residual ϵ_{ijt} to account for any transitory error. Using a two-way fixed effects model, we then regress R_{ijt} onto the fatality rate, person effects, and establishment effects.

$$R_{ijt} = \tilde{\gamma} a_{ijt} + \theta_i + \Psi_{J(i,t)} + \varepsilon_{ijt} \tag{3}$$

where θ_i is the person effect, $\Psi_{J(i,t)}$ is the effect of establishment J at which worker i is employed at time t. This identification strategy effectively decomposes the establishment effect into the component that affects wages through the fatality rate and the component that affects wages through all other unobserved channels, potentially including, for example, establishment productivity and compensation policies. The two-way fixed effects model, as in Abowd et al. (2002), is solved using a conjugate gradient algorithm.

The limitation of this model, however, is that it requires the restriction that the estimated match effect be orthogonal to both the worker and the establishment effect. While this a significantly weaker assumption than that imposed by any other VSL study to date, we later relax this restriction using an estimation approach that builds on Abowd and Schmutte (2013). The difficulty of this extension is that the full set of worker, establishment, and match effects are not identifiable using a fixed effects model. The alternative approach of using random effects is possible, but requires imposing the unpalletable assumption of orthogonality between the fatality rate and each of the latent effects. This assumption does not seem plausible. For example if workers unobserved heterogeneity affects their earnings and job choices and is also correlated with whether they choose to accept a safe or risky job, then this assumption does not hold.

5 Results

Table 3 compares estimates of the compensating wage differentials and implied VSLs using several identification stragies that are common in the literature, as well as our orthogonal match effects model. Model 1 is the traditional cross-sectional estimate. The estimated coefficient implies a VSL of about 670,000 Brazilian Reals (in 2003 Reals), conditioning on a cubic in experience, job tenure, gender, plant size, education, race, year, and state. In Model 2, which is a fixed effects specification that includes person effects, the estimated compensating differential falls by about 83% compared to the cross-sectional estimate. This upward bias in cross-sectional models compared to within-worker estimates has been well-documented in the literature using US data, including Brown (1980), Kniesner et al. (2012), and Lavetti (2014), and we find this same pattern in the Brazilian data. We also estimate Model 3, which has plant effects, but no worker effects. Identification in this model comes from two sources: variation over time in average industry-occupation fatality rates, and also from changes in the relative composition of jobs across industry-occupation cells within a plant over time. The former source is comparable to the type of variation used by Kniesner et al. (2012) in their analysis of workers who never change jobs, while the latter is comparable to cross-sectional variation in the composition of jobs at the firm. The estimate is even larger than the cross-sectional estimate, although it likely biased upward due to omitted worker heterogeneity.

The main improvement in the estimation relative to the previous literature comes from our estimates of the match effects models. Again, the identification comes not just from the often small and potentially non-salient changes over time in average fatality rates, but uses variation from workers switching jobs, and removes from the wage change associated with a job transition the effects of the change in unobserved plant and pure match effects, which would otherwise bias the estimated compensating differential. Qualitatively similar to the results of the match effects model in Lavetti (2014), we find that the estimate is substantially larger than the within-worker estimate (about three times larger), but not as large and the cross-sectional estimate. This implies that, although omitted heterogeneity from worker characteristics causes a large positive bias, conditional on unobserved worker effects, the residual plant and match heterogeneity actually causes a downward bias in the estimate. As a result, the within-worker model actually over-corrects for all sources of unobserved heterogeneity, and is potentially about as far from the unbiased estimate as the cross-sectional model is. As discussed above, however, this match effects model has the limitation that the estimated match effect is orthogonal by construction to the worker and plant effects. To the extent that this orthogonality assumption is invalid, this model could still be biased, although it unequivocally has fewer assumptions than the fixed effects models that have previously been estimated in the literature using PSID and other large data sources.

The remaining table repeat the same identification strategies, documenting heterogeneity in the estimates and in the relatives sizes of the biases for different subgroups of the Brazilian labor market. Table 4 shows results for men in full-time jobs. The cross-sectional estimate is substantially larger than that in Table 3, but after controlling work worker, plant, and match effects each of the fixed effects estimates is about the same as the comparable estimate using all workers. Two potential explanations for this are that the variance of unobserved heterogeneity among full-time men is larger than it is among other workers, or that the person effects are more strongly correlated with fatality rates.

Table **??** shows that the results are relatively insensitive to age. When excluding workers who are not of prime age (which we define as 23-65), the results remain largely the same.

Table 6 documents some interesting differences in estimation biases across genders. Although for men the omission of person effects biases estimates upward, for women the opposite is true. The female bias pattern is more consistent with the intuition that workers with higher ability spend some of their earnings capacity on safety and other occupational amenities, which, roughly speaking, is likely to lead to an increase in the estimated compensating differential when worker heterogeneity is removed. On alternative explanation that is consistent with findings in the behavioral literature, such as Barber and Odean (2001), is that men are overconfident about their own personal risk of death at work, whereas women are less likely to be overconfident. A gender difference in subjective expectations about risk levels could lead to a realized difference in gender-specific compensating wage differentials if women and men are imperfect substitutes for some jobs. Interestingly, the pattern of biases from unobserved plant and match effects appears to be the same for women and men, and it is only the bias caused by unobserved person heterogeneity that is markedly different.

For the sake of comparison to other studies in the literature, which have often used manufacturing workers as a focus, we show comparable estimates for male workers in manufacturing and production occupations in Table 7. The results are qualitatively similar to the results for all prime-age men.

Finally, we estimate the models similar to those in Table **??**, except that we include an orthogonalized cubic function of the level of risk. The orthogonalization procedure creates a cubic function that has the exact same variation as a standard cubic function, except that each term in the polynomial is constructed to be orthogonal to the lower-ordered terms. As a result the coefficient estimates are both more stable to specification assumptions, and the t-statistics can be easily interpreted as a test of whether the marginal contribution of each term is significant, conditional on all preceding terms. The cross-sectional parameter estimates suggest that the cubic model produces a significantly better fit than the linear model. However in the worker effects model the significance of the quadratic and cubic terms disappears for men but remains strong for women. The standard error estimates in the match effects model are not yet available, but will ultimately be more informative of the importance of allowing the fatality rate to have a non-linear effect on wages.

The importance of allowing for non-linearity in the model specification has been discussed previously, including by Ekeland et al. (2004), who argue that the assumption that the hedonic envelope function is linear is "arbitrary and misleading". However, the goal of the match effects model is to change the estimand entirely. Rather than estimating the hedonic equilibrium function that is characterized by the set of tangencies between workers' indifference curves and firms' isoprofit functions in wage-risk space, we instead seek to estimate the more primitive parameters that describe workers' marginal willingness to accept fatal risk. As shown by Hwang et al. (1998), the failure to account for endogenous mobility and non-random assignment of workers to firms causes the marginal willingness to accept to differ greatly from the hedonic equilibrium function.

6 Bias from Aggregation and Omitted Employer Characteristics [INCOMPLETE]

Our results indicate that panel estimates of the price of risk are contaminated by omitting employer and match specific heterogeneity. In this section, we develop an econometric framework to characterize different sources of omitted variable and aggregation bias. Our approach is related to the method for decomposing bias in industry wage premia discussed in Abowd and Kramarz (1999) and implemented in Abowd et al. (2012).

Consider a variation on Equation 2 in the spirit of the canonical two-way decomposition of earnings heterogeneity considered by Abowd et al. (1999):

$$w_{it} = x_{it}\beta + \theta_i + \psi_{G(i,t)} + \varepsilon. \tag{2}$$

The key difference in this specification is that now a unique "employer", g, is defined by a combination of plant and occupation. $G(i,t) = g$ if worker i was employed in plant-occupation combination g in year t. As before, w_{it} is the log wage, and θ_i captures characteristics of the individual that do not change over time and are correlated with wages.

Since job risk is a characteristic of the industry-occupation pair, the effects of job risk on earnings are absorbed by ψ_g in Equation 2. We define the true effect of risk on wages as the part of variation in employer pay associated with variation in risk. Consider the linear projection of the employer effects onto job risk:

$$\psi_g = \gamma a_{k(g)} + \phi_g, \tag{3}$$

where $k(g)$ aggregates the plant-occupation g to the industry-occupation level. This is an employer-level model. For now, we assume ϕ is uncorrelated with a, so in principle γ can be consistently estimated in an employer-level regression of plant-occupation effects from the first stage onto job-specific risk. We will relax this assumption in the empirical work. Let the number of observed plant-occupation pairs be G and the number of unique industry-occupation pairs be K. In matrix notation, the model above is expressed as

$$\psi = \gamma \Pi a + \phi, \tag{4}$$

where the $G \times K$ matrix Π classifies each of the plant-occupation pairs into one of the K industry-occupation pairs. a is the $K \times 1$ vector of observed fatality risks.[11]

The least-squares estimator for the true compensating wage differential, γ, is

$$\gamma = (a'\Pi'\Pi a)^{-1} a'\Pi'\psi. \tag{5}$$

We can insert pure compensating differentials into Equation 2 tautologically

$$w = X\beta + D\theta + \gamma F \Pi a + F\left(\psi - \gamma \Pi a\right) + \varepsilon. \tag{6}$$

[11] In this section, we assume for simplicity that the fatality risk does not change over time. Our results are not sensitive to accounting formally for the time-series variation in fatality risk.

The matrix notation follows Abowd et al. (1999): D is the design matrix of worker effects, and F is the design of employer (plant-occupation) effects. From the construction of γ it follows that $\psi - \gamma\Pi a = M_{\Pi a}\psi$ where $M_Z = (I - Z(Z'Z)^{-1}Z')$ is the idempotent 'residual-maker' matrix that projects into the column null space of Z.

6.1 Aggregation Bias

The magnitude of the compensating differential will be biased if we estimate it using job-level rather than employer-level data. Better jobs – higher paying and lower risk – are in larger plants and tend to last longer. If so, the relationship between risk and wages may be attenuated in job-spell data. Consider the employment-weighted version of Equation 4:

$$F\psi = \gamma F\Pi a + F\phi. \tag{7}$$

Even if ϕ is uncorrelated with Πa, it does not follow that $F\phi$ is uncorrelated with $F\Pi a$ since the number and duration of jobs of a particular type depend directly on risk.

It will be useful to define the duration and size-biased estimate of the compensating wage differential as

$$\tilde{\gamma} = \left(a'\Pi'F'F\Pi a\right)^{-1} a'\Pi'F'F\psi. \tag{8}$$

The duration bias is readily computed as

$$\tilde{\gamma} - \gamma = \left(a'\Pi'F'F\Pi a\right)^{-1} a'\Pi'F'F\psi - (a'\Pi'\Pi a)^{-1}a'\Pi'\psi, \tag{9}$$

implying again that the bias will be negative when employment is weighted toward high-wage/low-risk jobs.

Hwang et al. (1998) make a related observation that the magnitude of bias depends on whether the wage differential is estimated in worker-level or employer-level data. In their model, firms with high pay and low risk attract and retain more workers, so an employment-weighted analysis biases the estimated relationship between wages and risk downward.

6.2 Bias Decompositions

We now consider the effect of omitting unobserved heterogeneity. Our interest centers on what happens as we move from the pooled model to the within-worker model, and then to the model that controls for worker and employer effects. Let us first consider the model that includes worker effects, but omits employer effects. The estimated model is

$$w = X\beta + D\theta + \gamma F\Pi a + \varepsilon. \tag{10}$$

Let γ^* be the estimate of γ under this misspecified model. Estimates of the compensating wage differential that exclude employer effects are equal to the true effect, γ, plus a bias that is equal to the covariance between residual risk (after conditioning on D and X) and (employment-weighted) residual wage variation (after conditioning on job risk).

$$\gamma^* = \gamma + \left(a'\Pi'F'M_{[DX]}F\Pi a\right)^{-1} a'\Pi'F'M_{[XD]}FM_{\Pi a}\psi. \tag{11}$$

The bias is negative if observed employment is biased toward workers with idiosyncratically low risk who work in firms with idiosyncratically high pay. Our Equation 11 is very similar to Equation (3.10) in Abowd and Kramarz (1999), but modified to account for employment-weighting. This bias is equal to zero only in the special case $F\Pi a$ is orthogonal to $FM_{\Pi a}$ given worker characteristics. The underlying economics suggest this is unlikely to be the case.

Now consider estimation excluding both employer and person effects. The estimated compensating wage differential, now denoted by γ^{**} will be equal to the true effect plus a bias term that depends on the correlation between residual risk (controlling for X) and both residual worker effects and (employment-weighted) residual employer effects.

$$\gamma^{**} = \gamma + \left(a'\Pi'F'M_X F\Pi a\right)^{-1} a'\Pi'FM_X \left(D\theta + FM_{\Pi a}\psi\right). \tag{12}$$

Using the fact that $FM_{\Pi a}\psi = M_{F\Pi a}\psi$, we can rewrite Equation 12 as a weighted sum of worker and employer effects

$$\gamma^{**} = \left(a'\Pi'F'M_X F\Pi a\right)^{-1} a'\Pi'FM_X D\theta + \left(a'\Pi'F'M_X F\Pi a\right)^{-1} a'\Pi'FM_X F\psi. \tag{13}$$

The raw compensating wage differential is a sum of the correlation between risk and worker effects and risk and employer effects, conditional on time-varying observables.

6.3 Empirical Implementation

To perform the bias decomposition, we first estimate the AKM decomposition in Equation 2, controling for unobserved worker heterogeneity, and allowing employer effects to vary from occupation to occupation. We restrict the sample to dominant jobs with at least 30 hours contracted per week. The sample is otherwise unrestricted. Our analysis covers $136,217,638$ matches between $65,898,640$ unique workers and $12,142,644$ plant-occupation pairs.

We fit Equation 2 by the exact solution method described in Abowd et al. (2002). [12]

[12] Separate identification of the worker and plant-occupation effects is only possible within connected groups. While there are a large number of disconnected plant-occupations, over 95 percent of observations are in the largest connected group. We report results for the entire sample, but the results are not sensitive to restricting the sample to the largest connected group.

Table 10 describes employment duration-weighted moments of components of the log wage decomposition. The table reports the mean, standard deviation, and correlation among the log wage, time-varying characteristics, person effects, plant-occupation effects, and the wage residual. For reference, the table also reports the fatality rate, which, we emphasize, was not included in estimation. The statistics are computed by merging log wage components back to the estimation file, so these statistics are weighted by total employment and employment duration.

Table 10 indicates the fatality rate is weakly correlated with log wage and its components. This is partially due to the non-linear relationship around zero fatality risk jobs. Employers with fatality risk equal to zero also pay much more on average than firms with positive fatality risk. However, after controlling for whether an employer offers jobs with positive fatality risk, the relationship between fatality risk and the employer effects turns positive.

6.4 The True Compensating Wage Differential and Duration Bias

For our remaining analysis, we restrict attention to the same five percent sample of workers that are the subject of our prior results and the employers on their dominant jobs. Our estimate of the true compensating wage differential based on Equation 5 is $\hat{\gamma} = 0.0005253$. This result comes from projecting the estimated plant-occupation effects onto the fatality risk. Note, as with our previous results, we restrict the analysis to employers with positive fatality risk. This estimate therefore masks a considerable wage penalty associated with working in any risky job. Employers with positive fatality risk pay, on average, 14 percent less than employers with zero fatality risk after controlling for observable and unobservable worker characteristics. Note further that this estimate of the compensating differential is very close to the result from our orthogonal match effect specification in Table **??**. If we estimate, instead, Equation 7, which is the duration-weighted version of Equation 4, we obtain an estimate of the compensating wage differential, $\hat{\hat{\gamma}} = .0000721$. This is an order of magnitude smaller than the unweighted estimate.

6.5 Bias Decomposition

We focus first on the bias associated with failure to control for heterogeneity in plant-specific compensation. In the true model, equation 6, wages depend on observable and unobservable worker characteristics, the fatality risk, and idiosyncratic aspects of employer pay. Omission of residual employer-specific attributes of pay can bias the estimated compensating wage differential if, given worker characteristics, risk is negatively correlated with non-risk features of pay.

Tables 3 through 7 already suggest there is a negative bias from failing to control for match and plant heterogeneity. Using our dominant job sample, we also find that the estimated compensating wage differential is negatively biased if we fail to control for employer heterogeneity. The estimated compensating wage differential controlling for worker effects only is $\hat{\gamma^*} = 0.0002233$. From Equation 11, the bias term is

$$Bias = \left(a'\Pi'F'M_{[DX]}F\Pi a\right)^{-1} a'\Pi'F'M_{[XD]}FM_{\Pi a}\psi, \tag{14}$$

which is the coefficient from a regression of residual employer effects (conditional on risk) on residual risk (conditional on observed and unobserved worker characteristics). Fitting this regression directly, we estimate $Bias = -0.0001646$. As expected, this bias is negative, though smaller in magnitude than the algebraic decomposition would suggest.

7 Correlated Random Effects Model

In progress.

8 Graphical Evidence on Residual Correlations

Figure reffig:plot1 shows the average residual from the two-way fixed effects model, which controls for worker and plant effects, by decile of the worker effect and plant effect distributions. The figure suggests that high wage workers at high wage firms tend to have positive residuals that are economically meaningful, about 3-4% of wages. This suggests that job-assignment is nonrandom insofar as pure match effects are correlated with included worker and plant effects, causing endogenous mobility bias. The graph also shows that low wage workers at high wage firms tend to have negative match effects, as do high wage workers and low wage firms, consistent with the form of assortative matching in West Germany documented by Card et al. (2013).

Figure reffig:plot2 shows the average residuals from the same two-way fixed effects model by decile of distributions of fatality rates and worker effects. The figure shows that high wage workers tend to find better matches, and is consistent with the idea that workers who are better at finding good matches tend to spend some of this additional income on job amenities like safety, so that low risk jobs have systematically positive residuals. Both of these figures demonstrate that even when accounting for unobserved worker and plant heterogeneity that may be arbitrarily correlated with risk, this is not enough to resolve the problem of endogenous job mobility, motivating our development of the fully correlated random match effects model.

Figure reffig:plot3 plots the average change in wages associated with job changes by decile of the plant effects distributions. It shows that workers moving from the first decile of the plant effects distribution to the tenth decile experience a 186% increase in wages. It also demonstrates the importance of endogenous job mobility by showing the asymmetries in wage effects associated with job transitions. For example, a worker who moves from a job in the second decile to one in the fourth decile experiences a 23 percentage point increase in log wages, while a worker moving in the opposite direction experiences an 11 percentage point wage decrease. This asymmetry suggests that in some parts of the wage distribution, job mobility is associated with an unmodeled selection process.

In progress.

9 Conclusion

As has long been suspected, endogenous mobility of workers across jobs with different risk severely affects estimates of compensating wage differentials. Standard panel data approaches that correct for unobserved worker characteristics understate the true wage differential. Initial investigations of the nature of this bias indicate it has two sources: one is in the non-random selection duration of low-risk high-wage jobs and the second is the correlation between wages, risk, and firm size. These results suggest a model of compensating differentials with costly search in the spirit of (Hwang et al. 1998) may provide a useful guide to further empirical work.

References

Abowd, J. M., Creecy, R. H. and Kramarz, F. (2002). Computing person and firm effects using linked longitudinal employer-employee data, *Technical Report TP-2002-06*, LEHD, U.S. Census Bureau.

Abowd, J. M. and Kramarz, F. (1999). The analysis of labor markets using matched employer-employee data, *in* O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, chapter 40, pp. 2629–2710.

Abowd, J. M., Kramarz, F., Lengerman, P., McKinney, K. and Roux, S. (2012). Persistent inter-industry wage differences: Rent-sharing and opportunity costs, *IZA Journal of Labor Economics* **1**(1): 1–7.

Abowd, J. M., Kramarz, F. and Margolis, D. N. (1999). High wage workers and high wage firms, *Econometrica* **67**(2): 251–333.

Abowd, J. M., McKinney, K. and Schmutte, I. M. (2010). How important is endogenous mobility for measuring employer and employee heterogeneity? mimeo.

Abowd, J. M. and Schmutte, I. M. (2013). Endogenous mobility, *Cornell University School of Industrial and Labor Relations* .

Bonhomme, S. and Jolivet, G. (2009). The pervasive absence of compensating differentials, *Journal of Applied Econometrics* **24**(5): 763–795.

Brown, C. (1980). Equalizing differences in the labor market, *The Quarterly Journal of Economics* **94**(1): 113–34.

Card, D., Heining, J. and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality, *The Quarterly Journal of Economics* **128**(3): 967–1015.

Ekeland, I., Heckman, J. J. and Nesheim, L. (2004). Identification and Estimation of Hedonic Models, *Journal of Political Economy* **112**(S1): S60–S109.

Gasparini, L. and Tornarolli, L. (2009). Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata, *Desarrollo Y Sociedad* **63**(1): 13–80.

Groshen, E. L. (1991). Sources of Intra-industry Wage Dispersion: How Much Do Employers Matter?, *The Quarterly Journal of Economics* **106**(3): 869–84.

Hwang, H.-s., Mortensen, D. T. and Reed, W. R. (1998). Hedonic Wages and Labor Market Search, *Journal of Labor Economics* **16**(4): 815–47.

Kniesner, T. J., Viscusi, W. K., Woock, C. and Ziliak, J. P. (2012). The Value of a Statistical Life: Evidence from Panel Data, *The Review of Economics and Statistics* **94**(1): 74–87.

Krueger, A. B. and Summers, L. H. (1988). Efficiency Wages and the Inter-industry Wage Structure, *Econometrica* **56**(2): 259–93.

Lavetti, K. (2014). Dynamic labor supply with a deadly catch.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition, *Journal of Political Economy* **82**(1): 34–55.

Thaler, R. H. and Rosen, S. (1976). The value of saving a life: Evidence from the labor market, *Household Production and Consumption*, National Bureau of Economic Research, Inc, pp. 265–302.

Tsai, W.-J., Liu, J.-T. and Hammitt, J. K. (2011). Aggregation biases in estimates of the value per statistical life: Evidence from longitudinal matched worker-firm data in taiwan, *Environmental & Resource Economics* **49**(3): 425–443.

Viscusi, W. K. (2004). The Value of Life: Estimates with Risks by Occupation and Industry, *Economic Inquiry* **42**(1): 29–48.

Viscusi, W. K. and Aldy, J. E. (2003). The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World, *Journal of Risk and Uncertainty* **27**(1): 5–76.

Woodcock, S. (2008). Wage differentials in the presence of unobserved worker, firm, and match heterogeneity, *Labour Economics* **15**: 771–793.

Table 1: Descriptive Statistics

	Mean	Standard Deviation	N
Male	0.61	0.49	14,295,548
Age	34.46	11.18	14,295,548
Race *branco* (white)	0.52	0.50	14,295,548
elementary	0.15	0.35	14,295,548
less than HS	0.09	0.28	14,295,548
High School	0.38	0.49	14,295,548
Some College	0.04	0.20	14,295,548
College +	0.14	0.35	14,295,548
Contracted Weekly Hours	41.11	6.39	14,295,548
Log Monthly Earnings	6.52	0.74	14,295,548
Log Hourly Wage	1.37	0.80	14,295,548
Annual Earnings (2003 Reales)	9632.51	16555.10	14,295,548
Months worked	8.83	3.96	14,295,548
Total experience (years)	18.00	11.83	14,295,548
Fatality Rate	4.94	11.00	14,295,548
Fatality Rate Positive	0.83	0.38	14,295,548

NOTE–

Table 2: Fatality Rates By Industry and Occupation, 2003–2010

	Year							
Panel A: Sector	2003	2004	2005	2006	2007	2008	2009	2010
Agriculture and Fishing	11.649	12.271	10.528	9.371	7.941	8.356	9.044	6.514
Mining	12.035	13.579	12.235	7.460	6.242	11.044	9.119	9.871
Manufacturing	5.937	6.517	5.531	4.748	5.111	4.301	4.087	4.254
Utilities	6.060	5.459	4.430	4.318	4.969	4.452	2.000	1.603
Construction	15.188	15.948	11.785	12.109	13.086	10.642	11.298	9.115
Trade and Repair	7.319	7.881	6.774	5.679	4.753	5.217	4.625	4.464
Food, Lodging, and Hospitality	6.216	5.869	5.296	4.373	3.799	3.693	3.820	5.097
Transportation, Storage, and Communication	13.030	15.077	15.575	13.989	13.156	14.567	12.322	12.551
Financial and Intermediary Services	1.975	0.871	0.341	1.217	0.872	0.271	1.066	1.017
Real Estate	6.109	5.951	4.479	4.123	3.461	3.978	3.315	3.400
Public Administration, Defense, and Public Security	0.883	0.953	0.914	0.625	0.869	0.761	0.851	0.815
Education	2.069	2.787	1.276	1.590	1.231	1.380	1.076	1.236
Health and Social Services	1.884	2.253	1.879	1.357	1.087	1.862	1.424	1.345
Other Social and Personal Services	4.986	5.272	4.222	3.390	2.639	3.254	3.607	3.532
Domestic Services	21.504	8.823	0.000	0.000	0.000	15.409	0.000	0.000
Panel B: Occupation								
Public Administration and Management	3.213	3.495	2.860	3.033	3.008	2.402	1.950	2.395
Professionals, Artists, and Scientists	1.174	2.033	1.112	1.216	1.166	0.952	0.820	0.935
Mid-level Technicians	2.926	3.422	3.075	2.243	1.822	2.329	1.927	1.919
Administrative Workers	2.222	2.256	1.760	1.781	1.428	1.593	1.604	1.780
Service Workers and Vendors	5.687	5.711	4.592	4.012	3.317	3.695	3.659	3.575
Agriculture Workers, Fishermen, Forestry Workers	10.467	10.917	9.921	8.356	7.520	7.810	7.134	5.588
Production I	12.361	13.273	11.976	10.517	10.542	10.193	9.635	9.342
Production II	5.464	6.920	5.235	4.762	5.617	4.162	4.837	3.620
Repair and Maintenence Workers	8.795	7.731	6.733	5.689	5.865	7.954	7.395	6.712
Panel C: Overall	5.910	6.330	5.412	4.766	4.534	4.604	4.321	4.202

NOTE–Fatality Rates are expressed as deaths per 100,000 full-time full-year jobs.
SOURCE–Authors' calculations from RAIS microdata.

Table 3: Fixed Effects VSL Estimates: All Workers

	Dependent Variable: ln($Wage$)							
	(1) Cross-Sectional		(2) Within Worker		(3) Within Plant		(4) Within Match	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Fatality Rate (3-Yr MA)	0.000912*	(0.000225)	0.000170*	(0.000020)	0.001310*	(0.000235)	0.000508	
Experience	0.029181*	(0.000110)			0.026349*	(0.000094)		
Experience Sq.	–0.045911*	(0.000478)			–0.048902*	(0.000420)		
Experience Cu.	0.001361*	(0.000062)			0.002301*	(0.000055)		
Job Tenure	0.022477*	(0.000053)			0.011180*	(0.000043)		
Male	0.232469*	(0.001081)			0.127798*	(0.000541)		
Small Plant (Fewer than 20)	–0.281239*	(0.000744)			0.018110*	(0.001323)		
Medium Plant (20 to 499)	–0.145609*	(0.000708)			0.018935*	(0.001092)		
High School	0.270384*	(0.000501)			0.158010*	(0.000404)		
Some College	0.664732*	(0.001199)			0.381762*	(0.001035)		
College Degree	1.244734*	(0.001052)			0.855561*	(0.000861)		
Advanced Degree	1.920235*	(0.005988)			1.202686*	(0.005561)		
Black	–0.086865*	(0.000718)			–0.037633*	(0.000659)		
White	0.011432*	(0.000413)			0.032083*	(0.000423)		
yr2006	0.043643*	(0.000642)	0.075326*	(0.000399)	0.052321*	(0.000527)		
yr2007	0.082312*	(0.000664)	0.142059*	(0.000401)	0.087526*	(0.000583)		
yr2008	0.099392*	(0.000672)	0.197894*	(0.000404)	0.111782*	(0.000628)		
yr2009	0.086300*	(0.000685)	0.219569*	(0.000413)	0.101327*	(0.000657)		
yr2010	0.096289*	(0.000673)	0.271939*	(0.000426)	0.114413*	(0.000684)		
Constant	0.407253*	(0.001029)	1.141431*	(0.000347)	0.556342*	(0.001754)		
N	11,845,697		11,845,697		11,845,697		11,845,697	
R-Sq	0.431		0.878		0.705		0.921	
Worker Effects	N		Y		N		Y	
Plant Effects	N		N		Y		Y	
Match Effects	N		N		N		Y	
VSL	0.67		0.13		0.97		0.37	
95% CI	[0.35, 1.00]		[0.10, 0.15]		[0.63, 1.31]			

Notes: * Significant at the .01 level. All models include year effects and state effects. Sample selection criteria: fatality rate greater than zero.

Table 4: Fixed Effects VSL Estimates: Full-Time Men

	Dependent Variable: $\ln(Wage)$							
	(1) Cross-Sectional		(2) Within Worker		(3) Within Plant		(4) Within Match	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Fatality Rate (3-Yr MA)	0.001738*	(0.000059)	0.000170*	(0.000021)	0.001711*	(0.000092)	0.000473	
Experience	0.037237*	(0.000122)			0.030244*	(0.000119)		
Experience Sq.	–0.066685*	(0.000560)			–0.056386*	(0.000525)		
Experience Cu.	0.002736*	(0.000073)			0.002831*	(0.000067)		
Job Tenure	0.022796*	(0.000054)			0.011777*	(0.000053)		
Small Plant (Fewer than 20)	–0.299952*	(0.000617)			0.021306*	(0.001583)		
Medium Plant (20 to 499)	–0.153387*	(0.000572)			0.021709*	(0.001317)		
High School	0.246168*	(0.000464)			0.129260*	(0.000487)		
Some College	0.666940*	(0.001609)			0.357944*	(0.001524)		
College Degree	1.255216*	(0.001405)			0.796320*	(0.001479)		
Advanced Degree	1.610129*	(0.011237)			1.064287*	(0.009906)		
Black	–0.076089*	(0.000827)			–0.030106*	(0.000789)		
White	0.012551*	(0.000484)			0.039655*	(0.000523)		
yr2006	0.039812*	(0.000788)	0.071821*	(0.000471)	0.049692*	(0.000623)		
yr2007	0.088910*	(0.000774)	0.137227*	(0.000476)	0.088442*	(0.000641)		
yr2008	0.108207*	(0.000758)	0.196328*	(0.000480)	0.114758*	(0.000656)		
yr2009	0.100666*	(0.000756)	0.215590*	(0.000493)	0.108367*	(0.000671)		
yr2010	0.118728*	(0.000740)	0.271569*	(0.000512)	0.129328*	(0.000687)		
Constant	0.554129*	(0.001144)	1.133468*	(0.000428)	0.624881*	(0.001610)		
N	7,053,029		7,053,029		7,053,029		7,053,029	
R-Sq	0.398		0.884		0.719		0.931	
Worker Effects	N		Y		N		Y	
Plant Effects	N		N		Y		Y	
Match Effects	N		N		N		Y	
VSL	1.27		0.12		1.25		0.35	
95% CI	[1.18, 1.35]		[0.09, 0.15]		[1.12, 1.38]			

Notes: * Significant at the .01 level. All models include year effects and state effects. Sample selection criteria: men only, jobs with 30 or more hours per week, fatality rate greater than zero.

Table 5: Comparison of VSL Estimates

	Dependent Variable: ln($Wage$)							
	(1) Cross-Sectional		(2) Within Worker		(3) Within Plant		(4) Within Match	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Fatality Rate (3-Yr MA)	0.001799*	(0.000067)	0.000035	(0.000021)	0.001737*	(0.000103)	0.000416*	(0.000036)
Experience	0.030213*	(0.000288)	0.108035*	(0.000682)	0.025677*	(0.000234)		
Experience Sq.	–0.038464*	(0.001280)	–0.232022*	(0.002123)	–0.039085*	(0.001004)		
Experience Cu.	–0.000708*	(0.000166)	0.020731*	(0.000261)	0.000820*	(0.000128)		
Job Tenure	0.024621*	(0.000061)	0.005115*	(0.000058)	0.012260*	(0.000061)		
Small Plant (Fewer than 20)	–0.318293*	(0.000697)			0.024400*	(0.001754)		
Medium Plant (20 to 499)	–0.157854*	(0.000641)			0.023335*	(0.001452)		
High School	0.260366*	(0.000543)			0.130759*	(0.000561)		
Some College	0.708551*	(0.001844)			0.378123*	(0.001738)		
College Degree	1.261860*	(0.001453)			0.791876*	(0.001536)		
Advanced Degree	1.618880*	(0.011297)			1.064980*	(0.010064)		
Black	–0.079520*	(0.000931)			–0.032914*	(0.000893)		
White	0.016111*	(0.000547)			0.043564*	(0.000594)		
yr2006	0.037687*	(0.000890)	0.033621*	(0.000620)	0.048848*	(0.000693)		
yr2007	0.086942*	(0.000875)	0.059165*	(0.000892)	0.087645*	(0.000714)		
yr2008	0.103453*	(0.000859)	0.077574*	(0.001220)	0.114239*	(0.000732)		
yr2009	0.095205*	(0.000855)	0.055777*	(0.001563)	0.107334*	(0.000748)		
yr2010	0.111466*	(0.000838)	0.069206*	(0.001908)	0.127726*	(0.000768)		
Constant	0.589113*	(0.002145)	–0.053814*	(0.008545)	0.669513*	(0.002274)		
N	5,964,617		5,964,617		5,964,617		5,964,617	
R-Sq	0.380		0.893		0.722		0.941	
Worker Effects	N		Y		N		Y	
Plant Effects	N		N		Y		Y	
Match Effects	N		N		N		Y	

Notes: * Significant at the .01 level. All models include year effects and state effects. Sample selection criteria: men between ages 23-65, jobs with 30 or more hours per week, fatality rate greater than zero.

Table 6: Fixed Effects VSL Estimates: Full-Time Prime-Age Women

	Dependent Variable: ln($Wage$)							
	(1) Cross-Sectional		(2) Within Worker		(3) Within Plant		(4) Within Match	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Fatality Rate (3-Yr MA)	–0.000135	(0.000102)	0.000444*	(0.000076)	0.001609*	(0.000172)	0.000708	
Experience	0.015003*	(0.000254)			0.016119*	(0.000254)		
Experience Sq.	–0.010056*	(0.001128)			–0.023590*	(0.001103)		
Experience Cu.	–0.001387*	(0.000148)			0.000127	(0.000142)		
Job Tenure	0.022834*	(0.000079)			0.009104*	(0.000084)		
Small Plant (Fewer than 20)	–0.236046*	(0.000795)			0.015938*	(0.002337)		
Medium Plant (20 to 499)	–0.135335*	(0.000727)			0.016154*	(0.001937)		
High School	0.303615*	(0.000651)			0.171077*	(0.000725)		
Some College	0.676845*	(0.001612)			0.371629*	(0.001638)		
College Degree	1.134429*	(0.001163)			0.737691*	(0.001254)		
Advanced Degree	1.673128*	(0.010418)			1.070104*	(0.009084)		
Black	–0.112992*	(0.001272)			–0.047888*	(0.001282)		
White	–0.010876*	(0.000694)			0.011389*	(0.000795)		
yr2006	0.049971*	(0.001129)	0.074717*	(0.000630)	0.060735*	(0.000884)		
yr2007	0.081828*	(0.001110)	0.136449*	(0.000642)	0.093923*	(0.000905)		
yr2008	0.096529*	(0.001081)	0.187126*	(0.000647)	0.121166*	(0.000911)		
yr2009	0.085504*	(0.001069)	0.209036*	(0.000664)	0.116280*	(0.000919)		
yr2010	0.092808*	(0.001041)	0.255943*	(0.000691)	0.129823*	(0.000927)		
Constant	0.492734*	(0.002149)	1.091895*	(0.000568)	0.634909*	(0.002474)		
N	3,567,041		3,567,041		3,567,041		3,567,041	
R-Sq	0.424		0.906		0.748		0.948	
Worker Effects	N		Y		N		Y	
Plant Effects	N		N		Y		Y	
Match Effects	N		N		N		Y	
VSL	-0.09		0.31		1.12		0.49	
95% CI	[-0.23, 0.05]		[0.21, 0.41]		[0.89, 1.36]			

Notes: * Significant at the .01 level. All models include year effects and state effects. Sample selection criteria: women between ages 23-65, jobs with 30 or more hours per week, fatality rate greater than zero.

Table 7: Fixed Effects VSL Estimates: Full-Time Prime-Age Men in Manufacturing and Production

	Dependent Variable: ln($Wage$)							
	(1) Cross-Sectional		(2) Within Worker		(3) Within Plant		(4) Within Match	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Fatality Rate (3-Yr MA)	0.001867*	(0.000115)	–0.000201*	(0.000034)	0.001592*	(0.000249)	0.000353	
Experience	0.024109*	(0.000340)			0.018410*	(0.000245)		
Experience Sq.	–0.033551*	(0.001450)			–0.028053*	(0.000976)		
Experience Cu.	0.000452	(0.000185)			0.001042*	(0.000119)		
Job Tenure	0.017529*	(0.000081)			0.009234*	(0.000081)		
Small Plant (Fewer than 20)	–0.366043*	(0.001158)			0.012909*	(0.002336)		
Medium Plant (20 to 499)	–0.220706*	(0.000958)			0.017427*	(0.001981)		
High School	0.162825*	(0.000746)			0.035097*	(0.000769)		
Some College	0.418131*	(0.005326)			0.115429*	(0.004177)		
College Degree	0.589796*	(0.007001)			0.176839*	(0.006314)		
Advanced Degree	0.217634*	(0.037505)			0.057568	(0.026526)		
Black	–0.037193*	(0.001283)			–0.015468*	(0.001214)		
White	0.027198*	(0.000733)			0.025250*	(0.000786)		
yr2006	0.033226*	(0.001224)	0.053357*	(0.000779)	0.041877*	(0.000936)		
yr2007	0.065857*	(0.001200)	0.108479*	(0.000799)	0.081461*	(0.001027)		
yr2008	0.093470*	(0.001184)	0.157801*	(0.000816)	0.108825*	(0.001144)		
yr2009	0.077981*	(0.001183)	0.160274*	(0.000838)	0.099575*	(0.001217)		
yr2010	0.094589*	(0.001169)	0.205882*	(0.000872)	0.126339*	(0.001346)		
Constant	0.839923*	(0.002880)	1.144329*	(0.000823)	0.800482*	(0.004472)		
N	2,288,130		2,288,130		2,288,130		2,288,130	
R-Sq	0.247		0.850		0.691		0.913	
Worker Effects	N		Y		N		Y	
Plant Effects	N		N		Y		Y	
Match Effects	N		N		N		Y	
VSL	1.32		-0.14		1.13		0.25	
95% CI	[1.16, 1.48]		[-0.19, -0.1]		[0.78, 1.48]			

Notes: * Significant at the .01 level. All models include year effects and state effects. Sample selection criteria: men between ages 23-65, jobs with 30 or more hours per week in manufacturing and production occupations, fatality rate greater than zero.

Table 8: Fixed Effects VSL Estimates: Full-Time Prime-Age Men, Cubic Fatality Rate

	Dependent Variable: ln($Wage$)							
	(1)		(2)		(3)		(4)	
	Cross-Sectional		Within Worker		Within Plant		Within Match	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Orthogonalized Fatality Rate (3-Yr MA)	0.023346*	(0.000350)	0.000788*	(0.000283)	0.027295*	(0.000406)	0.008137*	(0.001023)
Orthogonalized Fatality Rate Sq.	–0.010479*	(0.000825)	0.000499	(0.000434)	–0.012757*	(0.000971)	0.002643	(0.003635)
Orthogonalized Fatality Rate Cu.	0.007348*	(0.000696)	0.001330	(0.000563)	0.008685*	(0.000823)	0.007811*	(0.002571)
Experience	0.030062*	(0.000287)	0.108022*	(0.000682)	0.025592*	(0.000233)		
Experience Sq.	–0.038130*	(0.001277)	–0.231986*	(0.002123)	–0.038997*	(0.001001)		
Experience Cu.	–0.000724*	(0.000166)	0.020728*	(0.000261)	0.000822*	(0.000127)		
Job Tenure	0.024745*	(0.000060)	0.005116*	(0.000058)	0.012269*	(0.000061)		
Small Plant (Fewer than 20)	–0.320729*	(0.000661)			0.024386*	(0.001753)		
Medium Plant (20 to 499)	–0.160252*	(0.000602)			0.023368*	(0.001451)		
High School	0.261752*	(0.000529)			0.131459*	(0.000559)		
Some College	0.711686*	(0.001823)			0.380081*	(0.001731)		
College Degree	1.265531*	(0.001418)			0.793374*	(0.001532)		
Advanced Degree	1.623223*	(0.011275)			1.067318*	(0.010043)		
Black	–0.080188*	(0.000928)			–0.032826*	(0.000892)		
White	0.015467*	(0.000543)			0.043581*	(0.000594)		
yr2006	0.038005*	(0.000889)	0.033650*	(0.000620)	0.049558*	(0.000691)		
yr2007	0.087825*	(0.000871)	0.059236*	(0.000893)	0.089373*	(0.000702)		
yr2008	0.104526*	(0.000854)	0.077669*	(0.001220)	0.116575*	(0.000710)		
yr2009	0.096385*	(0.000848)	0.055881*	(0.001564)	0.109931*	(0.000721)		
yr2010	0.112726*	(0.000830)	0.069324*	(0.001909)	0.130715*	(0.000732)		
Constant	0.605246*	(0.002091)	–0.053477*	(0.008546)	0.683115*	(0.002056)		
N	5,964,617		5,964,617		5,964,617		5,964,617	
R-Sq	0.381		0.893		0.722		0.941	
Worker Effects	N		Y		N		Y	
Plant Effects	N		N		Y		Y	
Match Effects	N		N		N		Y	

Notes: * Significant at the .01 level. All models include year effects and state effects. Fatality rate polynomial is orthogonalized using the Gram-Schmidt procedure. Sample selection criteria: men between ages 23-65, jobs with 30 or more hours per week, fatality rate greater than zero.

Table 9: Fixed Effects VSL Estimates: Full-Time Prime-Age Women, Cubic Fatality Rate

	Dependent Variable: ln($Wage$)							
	(1) Cross-Sectional		(2) Within Worker		(3) Within Plant		(4) Within Match	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Orthogonalized Fatality Rate (3-Yr MA)	–0.001633*	(0.000169)	3.136103*	(0.816501)	7.730433*	(1.390828)	0.014559	
Orthogonalized Fatality Rate Sq.	0.003208*	(0.000339)	1.509034*	(0.393578)	3.715963*	(0.670472)	0.002136	
Orthogonalized Fatality Rate Cu.	–0.003064*	(0.000692)	0.047094*	(0.012021)	0.117592*	(0.020505)	0.001897	
Experience	0.015056*	(0.000254)			0.016096*	(0.000254)		
Experience Sq.	–0.010268*	(0.001127)			–0.023533*	(0.001104)		
Experience Cu.	–0.001367*	(0.000148)			0.000120	(0.000142)		
Job Tenure	0.022762*	(0.000079)			0.009108*	(0.000084)		
Small Plant (Fewer than 20)	–0.234859*	(0.000792)			0.015982*	(0.002337)		
Medium Plant (20 to 499)	–0.134187*	(0.000724)			0.016197*	(0.001937)		
High School	0.302638*	(0.000649)			0.171443*	(0.000724)		
Some College	0.675759*	(0.001611)			0.372079*	(0.001637)		
College Degree	1.133107*	(0.001162)			0.737924*	(0.001254)		
Advanced Degree	1.672386*	(0.010422)			1.070295*	(0.009084)		
Black	–0.112214*	(0.001270)			–0.047849*	(0.001282)		
White	–0.010334*	(0.000692)			0.011499*	(0.000795)		
yr2006	0.049783*	(0.001129)	0.074793*	(0.000630)	0.061030*	(0.000884)		
yr2007	0.081303*	(0.001110)	0.136632*	(0.000643)	0.094603*	(0.000902)		
yr2008	0.095816*	(0.001081)	0.187364*	(0.000649)	0.122029*	(0.000907)		
yr2009	0.084720*	(0.001069)	0.209296*	(0.000666)	0.117196*	(0.000913)		
yr2010	0.092062*	(0.001041)	0.256202*	(0.000692)	0.130728*	(0.000922)		
Constant	0.492983*	(0.002123)	1.092910*	(0.000512)	0.638513*	(0.002404)		
N	3,567,041		3,567,041		3,567,041		3,567,041	
R-Sq	0.424		0.906		0.748		0.948	
Worker Effects	N		Y		N		Y	
Plant Effects	N		N		Y		Y	
Match Effects	N		N		N		Y	

Notes: * Significant at the .01 level. All models include year effects and state effects. Fatality rate polynomial is orthogonalized using the Gram-Schmidt procedure. Sample selection criteria: women between ages 23-65, jobs with 30 or more hours per week, fatality rate greater than zero.

Table 10: Correlation Among Components of the Log Wage Rate: RAIS 2003-2010

	Mean	Std. Dev.	Correlation					
			Log Wage	$X\beta$	θ	ψ	ε	Πa
Log Wage	1.30	0.760	1					
Time-varying characteristics	1.30	0.377	0.243	1				
Worker effect	−0.00	0.502	0.599	−0.476	1			
plant-occup. effect	−0.00	0.397	0.800	0.118	0.333	1		
Residual	0.00	0.196	0.258	−0.000	0.000	0.000	1	
Fatality Rate	5.28	10.594	−0.063	0.042	−0.095	−0.041	−0.000	1

NOTE–Correlation among components from the decomposition of log earnings into observable characteristics ($X\beta$), unobservable worker heterogeneity (θ), and unobservable plant-occupation heterogeneity (ψ) according to Equation 2. The column headers use symbols from the text while row headers provide short definitions. The fatality rate is not included in estimation, but reported here for comparison.
SOURCE–Authors' calculations based on RAIS microdata.

Figure 1: Mean residual by deciles of worker effect and plant effect

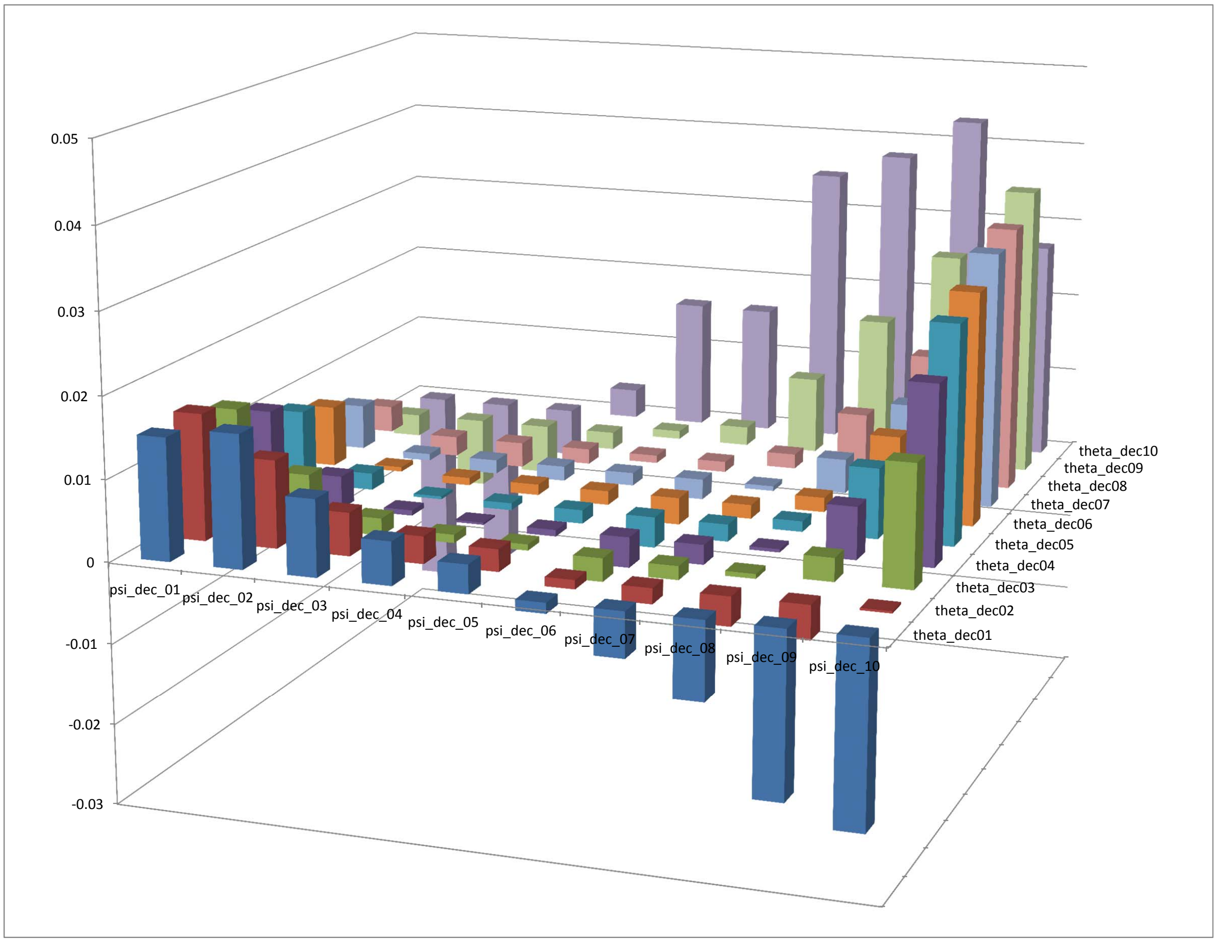


Figure 2: Mean residual by deciles of worker effect and fatality risk

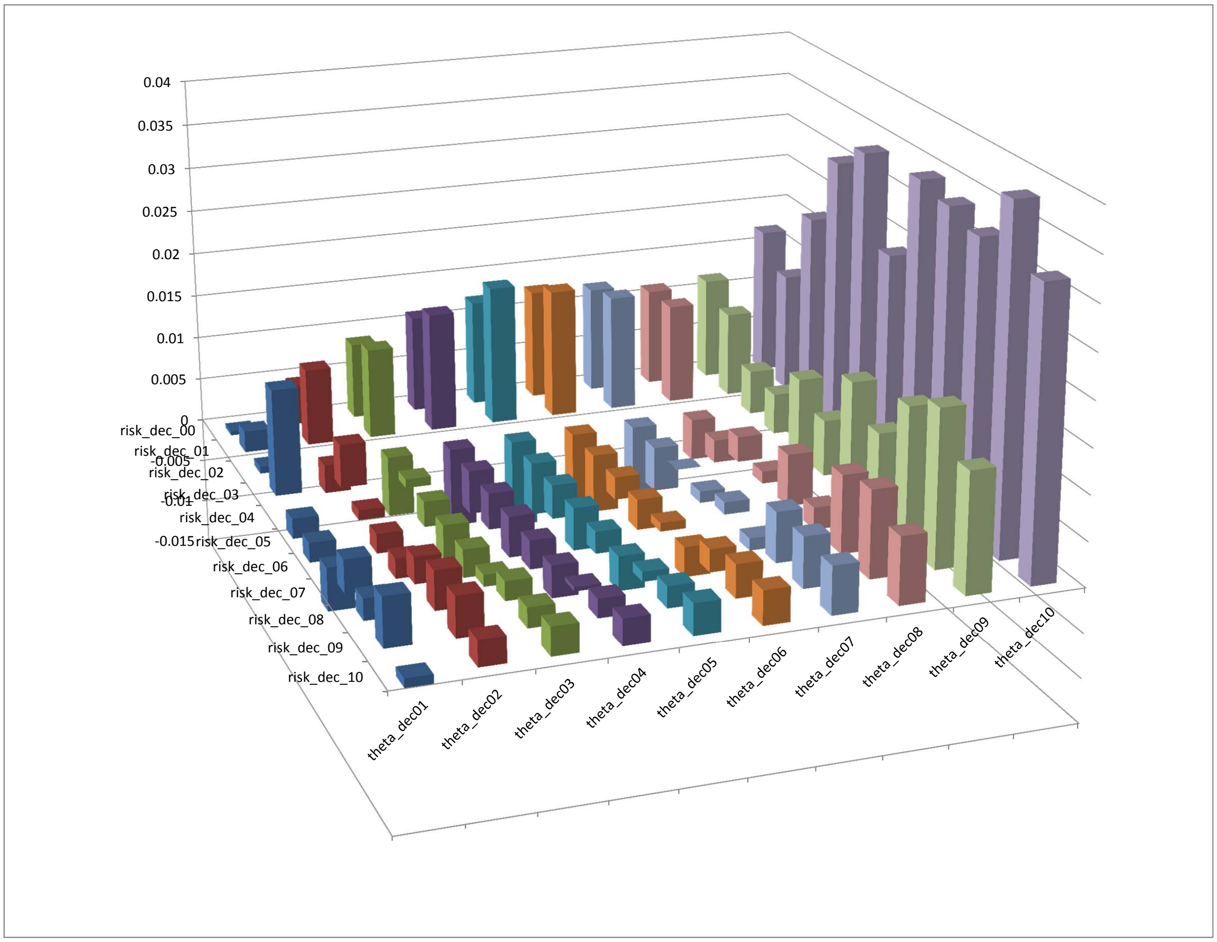
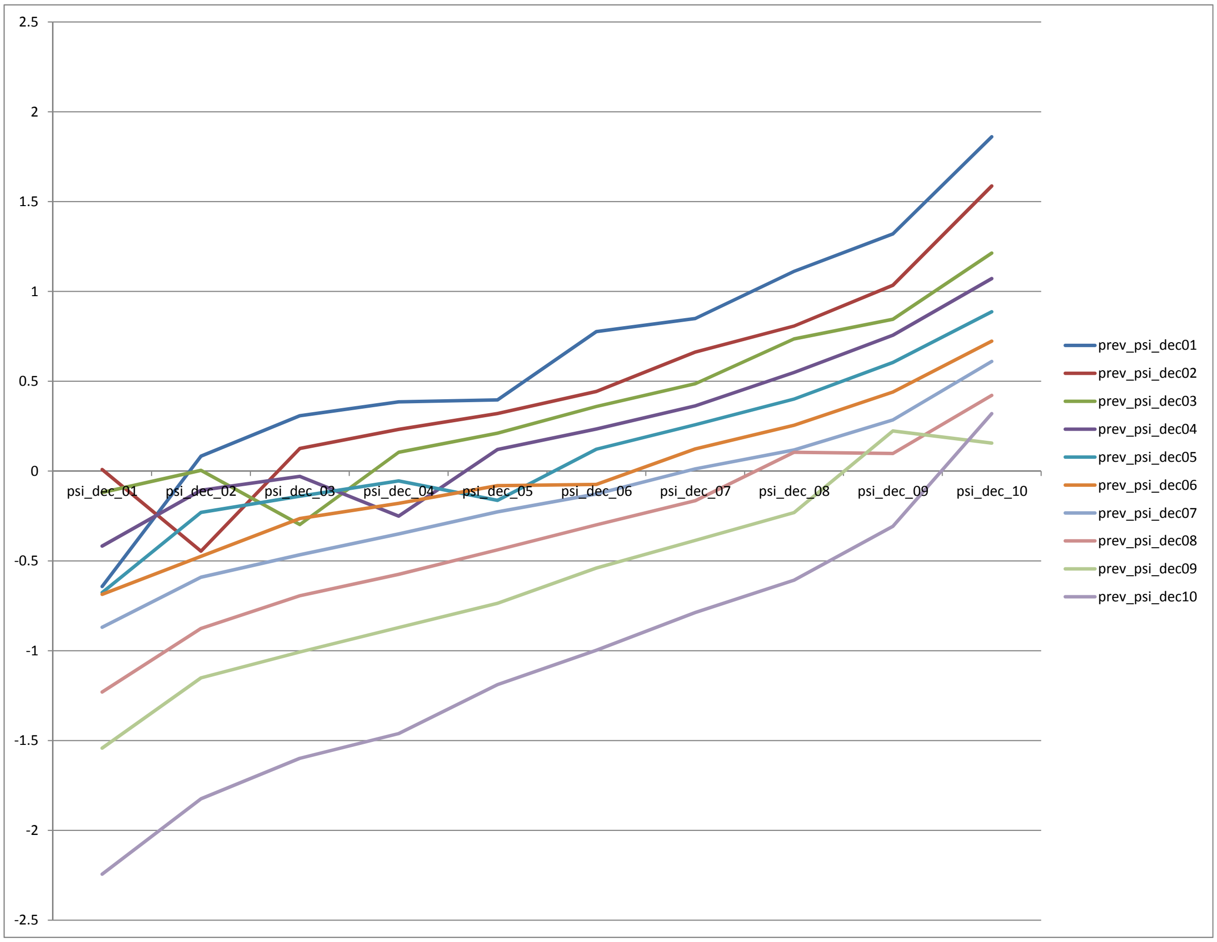


Figure 3: Mean difference in wages for workers by decile of the plant effect distribution



A Additional Tables

Table A.1: Causes of Separation Reported in RAIS

Value	Label Portuguese	Label English
0	nao desl ano	no separation this year
10	dem com jc	terminated with just cause
11	dem sem jc	terminated without just cause
12	term contr	end of contract
20	desl com jc	resigned with just cause
21	desl sem jc	resigned without just cause
30	trans c/onus	xfer with cost to firm
31	trans s/onus	xfer with cost to worker
40	mud. regime	Change of labor regime
50	reforma	military reform - paid reserves
60	falecimento	demise, death
62	falec ac trb	death - at work accident
63	falec ac tip	death - at work accident corp
64	falec d prof	death - work related illness
70	apos ts cres	retirement - length of service with contract termination
71	apos ts sres	retirement - length of service without contract termination
72	apos id cres	retirement - age with contract termination
73	apos in acid	retirement - disability from work accident
74	apos in doen	retirement - disability from work illness
75	apos compuls	retirement - mandatory
76	apos in outr	retirement - other disability
78	apos id sres	retirement - age without contract termination
79	apos esp cre	retirement - special with contract termination
80	apos esp sre	retirement - special without contract termination

NOTE–