Title

Employee Replacement Costs

Authors

Dube, Arindrajit
Freeman, Eric
Reich, Michael

Publication Date

2010-03-04

Peer reviewed

Employee Replacement Costs

Arindrajit Dube, Eric Freeman and Michael Reich[1]
Center on Wage and Employment Dynamics
Institute for Research on Labor and Employment
University of California, Berkeley

IRLE Working Paper 201-10
March 4, 2010

Abstract

We investigate properties of employee replacement costs, using a panel survey of California businesses in 2003 and 2008. We establish that replacement costs are substantial relative to annual wages and that they are associated negatively with the use of seniority in promotion. We also find some evidence, albeit not under all specifications, that replacement costs are positively associated with establishment size, which is consistent with monopsony. Bivariate scatterplots, pooled regressions and panel-based estimates suggest a positive relationship between replacement costs and the wage. While this result is not robust, it constitutes a puzzle for hiring and separation models, such as Manning (2003). In these models, the negative wage elasticity of replacement costs is a key assumption. These results thus call for further research on employment costs models.

[1]We thank Suresh Naidu for suggesting the study of the labor cost function and for helpful comments, Sylvia Allegretto for collaboration in constructing the second survey instrument, Steven Raphael for help with both instruments, and Sarah Frank for excellent research assistance.

1 Introduction

Although wages constitute a core topic in empirical and theoretical labor economics research, other labor costs to firms, such as the costs of replacing a worker, have been much less investigated. The cost of replacing a worker includes recruitment, selection and screening costs, the costs of learning on the job, and separation costs.[2] Indeed, microdata on the total cost of replacing a worker are surprisingly uncommon. A recent panel survey of California workplaces, the California Establishment Survey (CES), asks respondents about numerous characteristics of the establishments, including these replacement costs.

Figure 1 reports the mean, across establishments, of the average replacement costs for all workers in the workplace, as well as for two occupations, professional/managerial workers and blue collar/manual labor workers. The costs of replacing a worker can be substantial. Across establishments, they average about $4,000 overall, about $2,000 for blue collar and manual labor workers, and as high as $7,000 for professional and managerial employees.[3]

The mean ratio of an establishment's average replacement costs per recruit to the average annual wage among all employees, taken over all establishments, is 0.09. These costs constitute a significant fraction of annual wage costs and presumably affect their personnel decisions. Replacement cost is measured per recruitment event, and thus for a given position, annual replacement costs can be higher or lower, depending upon how often the average position turns over.

Data on recruiting and replacement costs are relatively rare, especially in combination with detailed workplace information. Thus, one focus of this paper is to report the levels and distributions of these replacement costs and their relationships with other variables. We also study in depth the properties of the replacement costs per employee, which is the total annual replacement costs at an establishment divided by the total number of workers.

The relationship of this quantity with the establishment's size is of particular interest. We find that replacement costs per employee grow with the size of the establishment, which indicates a monopsonistic labor market. We also find some evidence for a positive partial correlation between replacement costs per employee and the wage, which is at odds with an important model due to Manning (2003).

The plan of the paper is as follows. We first describe the California Establishment Survey and some of its key variables. We review the literature on replacement costs in section 3. Section 4 reports generally on replacement costs per recruit. Here we also investigate the association between replacement costs and the role of internal labor markets. We turn to replacement costs per employee and their relationship with establishment size, as well as the

[2]See below for details about the survey questions pertaining to replacement costs.

[3]Figure 1 also gives a sense of the precision of the estimates: the line segment at the right end of each bar represents a 95 percent confidence interval for the mean.

wage, in section 5. In section 6, we consider this latter relationship in light of the model of Manning and offer some comments regarding the apparent discrepancy. The final section concludes.

2 Data

The data set we use is a panel survey called the California Establishment Survey or CES (Survey Research Center, 2003 and 2008). The survey is a stratified sample of 1,080 establishments in California, conducted over the period May 13 to October 22 of 2003, and a follow-up survey of 652 establishments in California, conducted over the period April 29, 2008 to January 15, 2009.

We refer throughout to the two waves as the 2003 and 2008 waves. The 2008 survey data contains 358 establishments from the first wave, i.e., the CES includes a balanced panel of that size. Workplaces not contained in the 2003 sample were contacted in 2008 to increase the 2008 sample size; 294 in total were added. The survey instruments for each wave overlap a great deal, but do have many differences. The surveys were conducted by the Survey Research Center at the University of California at Berkeley.[4]

The CES includes data on numerous characteristics of each sampled establishment, including number of employees, pay scale, occupational distribution, employee tenure, hiring and training practices, general workplace practices, and benefits, as well as detailed information on the most recently hired non-exempt and professional employees.[5]

The sampling universe is the set of all nonprofit and for profit establishments in the state of California that have at least five employees, excluding public schools or universities, agriculture, forestry and fishing enterprises, as well as any government agencies. The sample was drawn from the California establishments database maintained by Dun and Bradstreet.

Sampling procedure

The CES used a stratified sampling scheme. The sampling process in 2003 was as follows: first, a stratified sample was taken of all establishments in California using a database maintained by Dun and Bradstreet (D&B). Larger establishments were selected at a higher rate. This sample (2806 in 2003) was checked to ensure it met the eligibility criteria, namely

[4] The Institute for Research on Labor and Employment at the University of California, Berkeley sponsored the survey. The instrument was developed primarily by Michael Reich for the 2003 wave and by Michael Reich and Sylvia Allegretto for the 2008 wave.

[5] The CES surveyed establishments. An advantage of surveying establishments, rather than firms, as pointed out by Osterman (1994), is that a respondent may be more in tune with the details of the workplace, and thus the data reported is likely to be more accurate. In the theoretical discussion throughout the paper, we may refer to either firms or establishments (workplaces), but in empirical work, we focus exclusively on establishments, as this is the level of observation in our data.

having at least 5 employees, and excluding government, public schools and universities, and agriculture, fishing and forestry; a check of eligibility was required as in some cases the D&B database was out of date or inaccurate. 2200 establishments met the eligibility criteria; of these, 1080 responded. In 2008, for the 1,080 establishments interviewed in 2003, an initial screen for still being in business and not having moved out of California reduced the sample to 1,016. Other eligibility criteria were then checked.[6] In addition to attempting repeatedly to contact all of the 2003 survey participants, an additional sample of new establishments was added to the 2008 survey to replace those that dropped out. This sampling process was similar to the original 2003 sampling process described above. Early response rates using the original 2008 survey instrument indicated that the length of the survey and the economic downturn were making it difficult to persuade workplaces to participate or to complete the survey. We therefore wrote an abbreviated version that omitted a number of questions from the original instrument.

Table 1 gives a summary of response rates for the sample. Establishments were split into 7 strata, based on Dun and Bradstreet's information on the establishment's number of employees. The groupings by number of employees for the strata were: 5-9, 10-19, 20-49, 50-99, 100-249, 250-999, and 1000+.[7] Larger establishments were oversampled, with, for example, in 2003, 100 percent of the largest stratum being sampled, while only 0.97 percent of the smallest establishments were sampled. Table 1 shows that the response rates ranged from about 35 to 50 percent, well within the normal range for establishment surveys. Response rates were somewhat lower among the largest workplaces (unreported).[8]

Finally, we briefly discuss our use of weights with the CES. All weights that we use account for the rate of oversampling by stratum, and account for nonresponse.[9] In some cases, one may want to adjust as well for the number of workers in an establishment, in case one wishes to consider questions where the worker is the natural unit of observation. We refer to these as worker weights, and weights without such adjustments as establishment weights. We use establishment weights in our calculations, unless mentioned otherwise.[10]

[6] In some cases, the 2008 D&B data for the establishments had missing data for the establishment size, or indicated that there were only 1 to 4 employees; the surveyors checked that all of these actually had at least 5 employees.

[7] Classification into strata used D&B data; in 2008 some establishments had no D&B data for establishment size or had a size of less than 5 employees in the D&B data. As mentioned earlier, the interviewers checked with the establishments that all of these cases actually had at least 5 employees. Generally, in the paper we use the 2003 strata for these establishments.

[8] See Survey Research Center (2003, p. 233, and 2009, p. 138).

[9] As is customary, we account for unit response, but not item nonresponse; that is, weights adjust for nonresponse at the observation level, but do not adjust for missing responses to individual questions.

[10] Note that percentages reported in Table 1 are simple unweighted percentages.

3 Related Literature

Data on recruitment and separation costs are hard to come by, but there is some work on this issue. Oi (1962) looked at 1951 data for a single company (International Harvester) and reported estimates of turnover costs as 7.3 percent of total labor costs for all employees and 4.1 percent for "common laborers." Manning (2006) reports that Campbell's work (1993), using a 1980 survey of employers, provides an implicit estimate of 8 percent of total labor costs.

Hamermesh (1993) reports some survey results.[11] A 1980 Los Angeles survey gave costs for hiring and training of $5,110 for production workers and $13,790 for salaried workers, which are $11,411 and $30,793, respectively, in 2003 dollars. A 1979 survey of large employers, at the national level, found that hiring cost $680 for secretaries and $2,200 for college graduates, which are $1,723 and $5,576, respectively, in 2003 dollars. On the other hand, in 1981, Merck & Co., a large pharmaceutical company, estimated the costs of turnover to be 1.5 to 2.5 times the annual salary of an employee, depending on the job. This study included, for example, a finding that in a new employee's first 14 months, essentially five and a half months of time was lost.[12]

Another more recent example (Abowd and Kramarz, 2003) examines the French labor market. Kramarz and Michaud (2010) expand on that work, reporting numerous findings, such as larger hiring costs for indefinite contracts than short-term contracts and greater costs for separation than for hiring. They estimate a firing cost of 34,983 French Francs per person in 1996, which is $8,358 in 2003 dollars, for workers fired either for cause or "for economic reasons;" and a hiring cost of 647 French Francs (155 dollars).[13] Manning (2003) reports data from the United Kingdom's Institute for Personnel Development on training and recruitment costs, and finds an average turnover cost of $\pounds 4,823$ for managerial and $\pounds 937$ for unskilled workers, or roughly $8,968 and $1,742, respectively, in 2003 dollars.[14]

Blatter, Mühlemann and Schenker (2008) look at two cross-sectional surveys of 4,032 Swiss firms on replacement costs, for the years 2002 and 2004. They look at the average hiring costs for workers with a vocational degree, and specifically include the costs of advertising the vacancy, interviewing expenditures and training costs. The average hiring costs over the two survey years are 13,500 Swiss francs, or roughly $8,300 in 2003 dollars, and these costs exhibit very large variation.[15] They also estimate the cost of hiring as a function of

[11] See Hamermesh (1993, p.208). In this section and throughout the paper we use the CPI-U to adjust for inflation.

[12] See Solomon (1988) for details on the Merck study.

[13] Here we use an exchange rate of 4.90850 Francs per dollar on January 1, 1998, accessed from http://www.oanda.com/currency/historical-rates .

[14] Here we use an exchange rate of .60710 dollars per British pound on January 1, 1998, accessed from http://www.oanda.com/currency/historical-rates .

[15] We use an exchange rate of 1.6583 Swiss francs per dollar on January 1, 2002, accessed from

the number of hires, and find that the marginal cost of hiring increases with the number of hires.

Regarding training costs, the American Society for Training and Development estimated the "average annual learning expenditure per employee" in 2008 to be $1,068 per employee. (Paradise, 2009) This figure includes internal costs, the cost of outsourced services and tuition reimbursement, which were 65.3 percent, 22.6 percent and 12.2 percent, respectively, of the costs in 2008. The average percentage of total annual payroll for these learning expenditures was 2.31 percent in 2003 and 2.24 percent in 2008.

Much of the literature focuses on adjustment costs and their functional form, and the manner in which adjustment costs affect how significantly a firm responds to demand shocks. We mention only a few examples of the work in this area. Hamermesh and Pfann (1996) discuss many of the central issues. Anderson (1993) uses data from the U.S. Unemployment Insurance system to show that adjustment costs can cause a significant dampening effect on firms' hiring and firing behavior in response to changes in demand.

The existence of replacement costs argues for search and matching frameworks that explicitly model these frictions, rather than the frictionless competitive model. (For an overview, see Mortensen and Pissarides, 1999.) Several recent matching papers (Mortensen and Nagypál, 2007, and Silva and Toledo, 2009) incorporate turnover costs as a key part of a strategy to better calibrate the model to empirical findings.

4 Replacement Costs per Recruit

In section 4.1, we describe the data on replacement costs available from the CES survey, and the summary measure we use, before reporting some descriptive statistics and basic graphs in section 4.2. We focus here on replacement costs per recruit, i.e. the costs of replacing one worker during one recruitment event.[16] In section 4.3, we investigate some relationships between replacement costs and other quantities.

4.1 Descriptive Summary Statistics

A question in the CES asked, for different occupational categories, the average cost of replacing a worker in that occupation. In the 2003 survey, this question includes "the cost of employee separation, recruitment, selection and screening, and on-the-job learning." In the 2008 wave, the question was the same except that "reading job applications" and "conducting interviews" were added.

http://www.oanda.com/currency/historical-rates .

[16]In section 5, we discuss replacement costs per employee, i.e. total costs for all recruiting divided by the total number of workers at the establishment.

Occupational categories differed between the two waves. In 2003, the occupational categories were: (i) professional or managerial employees; (ii) clerical employees; (iii) sales employees; and (iv) manual labor or blue collar employees. In 2008, the clerical and sales categories were combined into one category.

We require a summary measure to investigate replacement costs at the establishment level. To convert from variables at the level of the occupational category into a variable at the establishment level we weight by the fraction of workers in each occupation at the establishment. We refer to this constructed measure as the *replacement costs per recruit.*

Table 2 provides descriptive statistics for both waves of the CES, for our key variables. The estimates use data pooled from both waves of the survey, with a total of 1,732 observations. Averages are calculated at the establishment level, so they are not weighted by the number of workers in each establishment.[17] The calculations account for the stratification, and also cluster at the establishment level, to account for establishments present in both waves.[18]

In Table 2, the first column gives the mean of the variable, while the second column gives the mean of the log of the variable.[19] The average replacement cost per recruit is $4,039. Note that the standard deviation is quite large, at $9,800. The average of the log of the replacement cost per recruit is 6.41.[20] The mean establishment employs 44 workers.

The CES does not directly ask about the number of recruits or hires in the last year. The first recruits variables in Table 2, computed using what we call the 1 year rate, estimates this number by multiplying the fraction of full-time workers who have been at the establishment for less than one year by the number of employees at the workplace.[21]

Here we assume that part-time workers enter the establishment at the same rate as full-time workers, although turnover among part-time workers is likely to be higher than among full-time workers. Also, we do not observe if an employee joined the establishment in the last year but then left before the survey; this leads to an underestimate of the number of recruits. Our estimate of recruits is thus likely an underestimate of its true value.

The second recruits variable estimates the number of recruits by multiplying the number

[17] Standard errors are in parentheses while standard deviations are in brackets.

[18] Note that the calculations of the mean are not affected by the stratified nature of the sample other than by the weights, but the standard errors are. Following convention, standard deviations are estimated in a very basic way, by assuming one has a simple random sample of the same size. Strata used are those for each wave, except that 2003 strata are used for those 2008 observations having missing strata or stratum with establishment size 1 to 4. (All strata in the latter category are, by design, repeat observations from the 2003 sample.)

[19] There is one exception, which we note shortly.

[20] For many of the log calculations in the table, we use the logarithm of the sum of the corresponding variable and one, in order to avoid problems with taking the logarithm of zero.

[21] This fraction is either reported directly by the respondent, or alternatively the respondent can report the actual number of full time workers with less than one year of tenure, in which case we impute the fraction by dividing by the difference between the number of workers and the reported number of part time workers.

of workers by one fifth of the fraction of full-time workers who have been employed at the establishment for less than five years. We refer to this as the 5 year measure of recruits.

This estimation method decreases the variability of the measure when we construct our measure of labor costs, especially because the 1 year measure of recruits has numerous zeros for those establishments without any new hires in the last year. While the mean number of recruits in the last year is 6.9, the 5 year measure is 4.38. The 5 year measure's standard deviation is also lower. As the 1 year measure is more variable than the 5 year measure, extreme values may be pulling up the mean; the logs are not very different, which supports this explanation.

The hourly wage is a summary figure for each establishment, estimated using responses to the fractions of workers in various wage brackets. Averaging across all workplaces gives a mean of $18.55 an hour.[22]

The entry wage is the entry wage for the establishment's most common occupational category. We report the fraction professional or managerial, and the fraction of blue collar or manual labor workers, but not the fractions for clerical and sales workers as these questions changed between the two waves. Other notable summary statistics estimates include: at the average establishment, 31 percent of workers have a college degree and 18 percent work part time, and 82 percent of workplaces offer health insurance plans to their employees.

Regarding union density, at the average establishment, only 3 percent of workers are covered by a collective bargaining agreement. When one weights by the number of workers, this percentage (not reported in the table) increases to a level much closer to that typically reported for private sector union coverage.[23] (Recall that the CES does not include government establishments.)

The CES includes numerous questions about training and recruiting. 19 percent of establishments have a department explicitly for training employees. The survey also asks about the number of months it took for the most recently hired employee to be fully productive, in the category of professionals and managers, and the category of non-professional, non-managerial workers.[24] Across establishments, the mean number of months to full productivity is 2.6 for professionals and managers, and a smaller 2.0 months for non-professional and non-managerial employees. A similar question asked for the time to full group produc-

[22] The figure for the logarithm is the one case in which the second column reports a quantity other than the mean of the log of the average; here it is actually the mean of a summary figure computed for each establishment, i.e. the mean, across establishments, of the average of the log wages at the establishment.

[23] Our worker-weighted estimates are 12.3 percent in 2003 and 8.8 percent in 2008.

[24] There are some issues of the exact comparability of these in the definition of non-professional across waves, and the questions were asked in a different way across waves. The question is asked for employees that are neither professional nor managerial in 2008, but for clerical and blue collar employees in 2003. In 2003, a question asked for the number of months to full productivity (in buckets), whereas in 2008, respondents were first asked if it took less than one month, and if not, were then asked the number of months (but directly, i.e., not in buckets).

tivity after a new hire, again for each of the occupational categories mentioned above.[25] Professional or managerial workers take about 3.3 months to reach full productivity whereas non-professional workers take about 2 months. 30 percent of establishments require a written test for some hires.

Concerning recruiting methods, 52 percent of workplaces recruit on the internet, while 29 percent post help-wanted signs. It is interesting to note the changes in these numbers. By looking at the observations in the balanced panel which have nonmissing data on the relevant question in both waves, we see that internet recruiting rose from 38 percent to 65 percent of establishments from 2003 to 2008, while use of help wanted signs remained stable, at 28 percent in both 2003 and 2008.[26]

4.2 Detailed Descriptive Statistics

In this section, we look at replacement costs (per recruit) in more detail. Figure 1 presented their means, pooled across the 2003 and 2008 surveys. Here we consider these costs separately for each wave. Figure 2 gives a summary measure for all workers at each establishment, by occupational group and by survey wave.[27] Table 3 reports the average replacement costs for each occupational category for the cross-sections in each year (denoted as 2003 or 2008 Cross Section), and also the replacement costs in each year for those establishments that are present in both waves of the surveys, which we denote as 2003 or 2008 Balanced Panel.

The measures for the balanced panel subsample allow us to compare the change in cost across years using the variation within establishments.[28] The results, with standard errors and standard deviations, are reported in Table 3.

The average replacement costs for all workers in the workplace are in the range of $3,000 to $4,500, depending on the sample. Generally, professional and managerial replacement costs are highest, followed by costs for sales workers, clerical workers and finally blue collar and manual labor workers. Blue collar replacement costs are on the order of one fifth to one quarter of costs for professional and managerial employees.

By examining the balanced panel columns in Table 3, we can compare costs for all workers, as well as for the two occupational categories of professional and managerial workers, and

[25] This was asked for a new hire in general, not necessarily the most recent hire. For full group productivity, there were similar differences across waves in how the questions were asked, and how categories are defined, as for the time to full productivity questions.

[26] Using the entire cross-section also shows an increase in internet recruiting, but a small drop in help wanted sign usage, from 31 percent to 26 percent.

[27] These results are means across establishments. The line segments at the top of each bar represents a 95 percent confidence interval for the mean.

[28] To better achieve this comparability in the case of the balanced panel, we also restrict, for each measure, to those observations that have nonmissing data for that particular measure *in both years.*

blue collar and manual labor employees.[29] We see an increase in (real) replacement costs for all workers, and for professional and managerial workers, but very little increase for blue collar employees.[30]

These results are at odds with the overall decrease in replacement costs across the cross sections. This could result from the small sample size, or from a higher nonresponse rate among large establishments, which generally have higher replacement costs, as we will see later.[31] The effect of a differential nonresponse rate should be mitigated because the sample is stratified by establishment size, although there could still be nonresponse bias within each stratum.

Figure 3 presents histograms of replacement costs per recruit, in order to highlight the long right tail of these costs. The top chart presents costs for all workers, a weighted average of each establishment's responses by occupational group. The bottom two charts present establishment responses for the two occupational groups that are comparable across waves.[32] Estimates of the densities use the sampling weights. Although there are outlying observations as large as $300,000, we focus only on the subset of costs less than $25,000.

The top graph in Figure 3 shows that most establishments have replacement costs (averaged over all employees) lower than $10,000. There is a fairly long right tail; 2.5 percent of establishments have replacement costs greater than the $25,000 cutoff and thus are not represented in the histogram. The graph for professional or managerial workers, in the bottom left, shows lower densities near 0 and a larger right tail; here 5.5 percent of establishments have average replacement costs for professionals and managers of more than $25,000. Finally, for blue collar and manual labor employees, the distribution is much more bunched near 0, with a large spike in the $0 to $500 bin, while only 0.5 percent of establishments have average replacement costs greater than $25,000 for this category of employees. The histograms demonstrate the large amount of variation in replacement costs in general and particularly the long right tail of the distribution; moreover, replacement costs for blue collar workers are unsurprisingly much more concentrated near $0 than costs for professional workers.

We examine next the ratio of replacement costs (per recruit) to the average annual wage for workers.[33] Figure 4 shows the mean and median of this establishment-level variable, by survey wave; the mean was reported earlier for the pooled sample.

[29] One might reasonably argue that these are not strictly comparable across survey waves because of the different categorizations of clerical and sales workers.

[30] The sample sizes are, respectively, 130, 161, and 163 for all workers, professional/managerial employees and blue collar employees.

[31] The nonresponse could be either at the level of the observation (i.e., unit nonresponse) or the question (i.e., item nonresponse).

[32] As for the other questions, clerical and sales workers were treated as distinct in 2003, but combined into one group in 2008.

[33] To be clear, we now return to discussion of all workers in the establishment, not simply particular occupations.

In Figure 4, 2003 and 2008 Cross Section again denote cross sections, and 2003 and 2008 Balanced Panel again denote the balanced panel subset for each year (the subset with complete observations of the variable in both waves). Replacement cost here is measured per recruit. The mean ranges from about 6 to 10 percent, while the median sits in the neighborhood of 2 to 4 percent. The median indicates that these costs are a sizable consideration for establishments when hiring or firing; the large size of the mean indicates that for many firms, the costs are a significant factor in labor force decisions.

We compare across years by examining the balanced panel subsample. We see a decrease in the mean of the ratio, yet a slight increase in the median, but any comparison should bear in mind the small sample size ($n = 124$). In summary, the graph demonstrates that replacement costs play an important role in personnel decisions.

We now look at some of the quantities that affect replacement costs: we focus on training hours for new employees and the time it takes for a new employee to reach full productivity.[34] Figure 5 presents a histogram of this quantity, using the pooled sample. Generally, responses are clustered in the zero to 40 hours range. A large fraction of respondents gave the even answers of 40 and 80 hours, resulting in spikes at these spots, and other multiples of 40, as well as at 100 and 300 hours, the latter due to the topcoding. The bins are of size 5, so the spike at 40 hours represents a fraction of roughly 20 percent of respondents, while on the order of 5 or 6 percent give a response of 300 or more training hours.

Now consider the number of months to reach full productivity. As discussed above, we have measures for professionals, as well as a combination measure for non-professionals.[35] In the top two graphs in Figure 6, we give histograms for the pooled sample, for each category.[36] Note that the $y-$axis here measures the fraction of observations, as opposed to the density. To improve readability, we do not show observations with responses greater than 12 months.[37] Relative to the professional histogram, the mass in the distribution is shifted to the left in the non-professional histogram. But, surprisingly, the plots do not look particularly different.

The bottom two plots compare across waves within the balanced panel, restricting to observations for which the relevant variable is present in both waves. A response of less than one month was offered as a response in the 2008 wave but not in the 2003 wave; we have represented these as .5, which explains the large difference between waves. However, when adding the two left-most values of each graph, one obtains similar values. In general, the two distributions are similar in terms of having more weight on smaller numbers of months,

[34]We topcode the training hours variable at 300 hours to limit the influence of exceptionally large values. The largest response is 40 hours a week for 3 years, which we impute at 6000 hours.

[35]Recall that we noted above that neither is strictly consistent across waves, but particularly the latter.

[36]Respondents report ranges of values; we represent these by the midpoints of the ranges.

[37]There are 3 such responses for professionals, and 2 for non-professionals.

but otherwise, they are hard to compare due to the different ranges used across waves in the questions.[38]

So, regarding quantities that comprise replacement costs, training hours for new employees are significant enough that establishments undoubtedly consider them in formulating labor policies, and can be quite large in some cases. The time for workers to reach full productivity ranges from less than a month to a year or more, and are thus also important. Although hindered by differences in question wording, the times to full productivity are fairly similar across waves.

4.3 Replacement Costs and Seniority in Promotion Decisions

There is reason to believe hiring and promotion policies have an important effect on replacement costs. We consider here whether one such policy, the use of seniority in promotion, is negatively correlated with replacement costs per recruit.[39] This is to be expected, for two reasons. For one, when replacing workers from within, prioritizing higher seniority could reduce training costs if senior workers are more likely to have establishment-specific skills. Alternatively, the interviewing and hiring process could be much less expensive if the simple criterion of seniority is used to make a decision, as opposed to a more prolonged process.

The CES asked, if an establishment has already decided to fill a position from within, how important seniority is in the decision about who gets the job.[40] The answer is coded on a four point scale from 1 to 4, ranging from slightly to extremely important, respectively. The mean across establishments, shown in Table 2, is 2.18.

To test the seniority hypothesis, we regress the 2003 to 2008 difference in the log of the average replacement costs at an establishment on the difference in the seniority importance variable. We also control for differences in other variables, and we include a constant in the regression, which allows for the possibility of an additive change in replacement costs.

Table 4 displays the results. In the first column, we have a bivariate regression with a constant term, to allow for a constant change across the waves. An increase of importance of 1 point up the four point scale from 2003 to 2008, e.g. from moderately to very important (2 to 3), is associated with roughly a decrease of 64 percent $(1 - \exp(-1.03))$ of the average cost of replacing one recruit, which is quite a strong effect. This estimate is statistically significant at the 0.1 percent level.[41]

[38] See notes for Figure 6 for details of the ranges.

[39] Fairris (2004) found a negative relationship between the importance of seniority in promotion decisions and the voluntary quit rate.

[40] The precise form of the question is “When a vacancy for a job is filled with a current employee, how important is it for you to use seniority as a decision criterion?”

[41] Note, however, the small sample size; the balanced panel itself has 358 observations and nonresponse for the replacement cost questions was sizable.

In the second column, we add controls for the log of establishment size and the average of the log wage. (These variables are the difference of their value in 2008 from their value in 2003.) The third column includes controls for the fraction of workers covered by a collective bargaining agreement, as well as the fraction in occupational categories, the fraction with a college degree and the fractions of temporary and part-time workers.

In the latter two columns, the number of observations is reduced, but the estimate of the importance of seniority remains about the same size, at about -0.91. The estimate of the coefficient of the average log wage is positive and almost statistically significant at conventional levels.

The estimate on the fraction of blue collar workers is surprisingly positive and statistically significant at the 5 percent level. We do not include a coefficient for whether or not an establishment offers a health insurance plan, as only 5 establishments had changes in this indicator from 2003 to 2008, and thus such a variable when differenced would be nearly always zero.

To check our results, we also looked at a regression of the log of replacement costs per recruit on the seniority importance variable, but on the pooled sample. In these results, which we do not report here, we find estimates that are similarly negative, both using bivariate regressions, and using multiple controls, including 1-digit industry controls. However, the point estimates are about one-third to one-half of the size, and two out of the four regressions give estimates that are significant only at the 7.4 percent and 9.5 percent levels. In these unreported regressions, the estimates for the fraction of blue collar workers, while still positive, are less than one fifth of the above estimates, and are not statistically significant at any conventional level.

Overall, we find strong evidence that prioritizing seniority in promotion is associated with reduced replacement costs.

Another hiring policy, the importance of giving preference to current workers when filling a job, does not appear to be associated with replacement costs (results again not reported). A possible interpretation is that workplace-specific training is only valuable to an establishment for those workers who are sufficiently senior to have acquired enough training and associated skills. Alternatively, the mechanism of seniority providing a simple promotion criterion may be the key reason for our seniority finding.

5 Replacement Costs per Employee

As mentioned above, the actual cost of turnover to a workplace depends not only on the replacement cost per instance of recruitment, but also on how often it needs to recruit for a particular position. To this end, we examine *replacement costs per employee,* which is simply the replacement *cost per recruit* multiplied by the ratio of the number of recruits to the number of employees at the establishment.[42] Alternatively, this is the product of the cost of replacing one recruit and the recruitment rate.

The relationship between replacement costs per employee and establishment size is of particular interest. If the marginal cost of replacing a worker increases as the establishment size increases, then, in other words, an employer finds it more difficult to hire workers as the workplace grows. Thus these frictions are characteristic of a monopsonistic labor market.[43]

Throughout the section, we consider how replacement costs per employee vary with many other quantities. One of these is the wage. The relationship is certainly interesting in and of itself, but it also plays a key role in a model we later investigate.

5.1 Variable Definitions

In our measure of labor costs, we use the measure of the number of recruits described above, which uses in its construction one fifth of the percentage of full time employees working at the establishment for fewer than 5 years. We multiply the responses given for the replacement costs per recruit by the ratio of the number of recruits to the number of employees.

We also construct two alternative measures of labor costs; we use these in unreported calculations to check the robustness of our results. For one, we do the same as above, but use survey responses about the number of employees working at the establishment for less than 1 year, which we refer to as the replacement costs per employee using the 1 year recruitment rate, as it uses the ratio of recruits in the last year to employees.

Finally, we construct a measure of replacement costs per recruit that uses data on the number of training hours for new workers and the time needed for workers to achieve full productivity in order to estimate replacement costs. This is combined with the 5 year recruitment rate to produce a measure of replacement costs per employee. We refer to this as the imputed replacement costs per employee.

Table 2 reports the means and standard errors of these variables and their logs.[44] The mean value of the replacement costs per employee is \$439, which is 1.2 percent of the mean annual wage, assuming a 40 hour workweek and a 50 week year. Using the 1 year rate (as opposed to the 5 year rate) to estimate recruits yields a much higher mean and a much larger standard deviation. The mean of our imputed replacement costs per employee are reassuringly close to the mean of our measure from interviewee's responses. These measures are discussed in more detail in Appendix *A*.

[42] We also refer to the quantity as replacement *cost* per employee.

[43] We return later to a more formal discussion of this idea.

[44] To be clear, these are the logs of the entire ratio of replacement costs per employee, here and throughout the paper; they are not the ratio of the log of the numerator to the log of the denominator.

5.2 Graphical Analysis

We turn now to an analysis of replacement costs per employee and their relationship with establishment size, starting with some simple bivariate plots.

Figure 7 displays a plot of the log of replacement costs per employee against the log of establishment size. We use log specifications here and in what follows for ease of interpretation as elasticities, and because plotting levels would bunch observations into one corner of the graph. Here each circle represents an observation, and each circle's area is proportional to the establishment's nonresponse-adjusted weight in the survey.

We also graph a local polynomial smoother (with bandwidth 0.5) on the same plot. Replacement costs appear to be correlated positively with establishment size, consistent with a monopsony model.[45] Of course, this is only a bivariate correlation, and it may be that we are omitting an important quantity, which, once controlled for, would lead to a negative correlation between replacement costs and the establishment size.

In Figure 8, we plot the log of replacement costs per employee against the log of the establishment's average log wage. As above, each circle represents an observation, and each circle's area is proportional to the establishment's survey weight. We again graph a local polynomial smoother (also using bandwidth 0.5) on the same plot. Replacement costs and the wage are also strongly positively correlated.

5.3 Regression Analysis

Table 5 presents regression estimates concerning the replacement cost per employee and its relationship with establishment size and other variables. In the first column, we regress the log of the ratio of replacement costs to employees on the log of establishment size. In the second column, we add the average log wage at the establishment as a regressor. In the third column, we add multiple controls, and in the final column, we include 1-digit industry controls. The regressions are pooled across both waves. In other words, the observations are at the establishment-wave level (so if an establishment appears in both waves, it is treated as two observations), and we cluster at the establishment level.

In specifications (1) and (2) the estimate of the coefficient on the log of the establishment size is positive and statistically significant, with a $p-$value of .001. This finding indicates support for the monopsony model. This result obtains despite a mechanical reason that might produce a negative association: we divide by the establishment size in the course of calculating replacement costs per employee. The estimate of the coefficient on the average

[45] Although it gives the same result in the right section of the graph, one should ignore the smoother in this region, as there are few points and thus the smoothing estimates are liable to have large associated standard errors.

log wage, in the second column, is positive and statistically significant with a $p-$value of less than .0005.

In specification (3), we add controls for the fraction of workers at the establishment covered by collective bargaining, the fraction in two occupational groups, the fraction having a college degree, the fraction of temporary employees, the fraction of part-time employees, and whether or not the establishment offers a health insurance plan. In specification (4), we keep all of these controls and add 1-digit industry fixed effects, dropping the constant term. In these final two columns, we still obtain positive and strongly statistically significant estimates of the coefficient of log establishment size, and also of the coefficient of the log wage. The health insurance plan is the most statistically significant other control here, at the 4 percent or 5 percent level, and has a positive association with replacement costs in both regressions, as one might expect because of the costs of administration of new employees' health plans.

When we check the robustness of these results using our two alternate measures of replacement costs per employee (not reported here), we find fairly similar results. The magnitude of the estimates for the coefficient on log establishment size vary, but are statistically significant at the 5 percent level in all specifications. However, when we use our imputed replacement costs, and after including controls, the association with the log wage is much smaller in magnitude, and is no longer statistically significant.

To investigate these relationships further, we make use of the panel nature of the data. Specifically, we estimate a regression of differences in the outcome variable, from 2003 to 2008, on differences in the regressors. This approach is equivalent to a regression with a fixed effect for each establishment, and thus would likely control for unobserved characteristics of each establishment.[46]

Table 6 presents the regression of changes on changes, using only the balanced panel. We include a constant, which allows for a constant shift in the log of replacement costs per employee across all establishments. In the first column, we regress on only the log establishment size; in the second column, we include the average log wage. The final column adds controls.[47] Variables here are the difference of their value in 2008 from their value in 2003 (except of course the constant).

Despite having 358 observations in the balanced panel, the sample size is fairly small: for example there are 117 observations used in the first regression. The reason is that there are numerous missing responses for the underlying variable replacement costs per recruit, and unsurprisingly it is fairly likely that an establishment will have missing data for this variable

[46]This approach would not capture all effects of unobserved characteristics, however, if the relationship between each establishment's labor costs and its size changes across waves in a way other than by a simple constant additive shift in replacement costs per employee across all workplaces.

[47]There are no 1-digit industry controls, as it is rare for an establishment to switch industry across years.

in at least one of the two waves.

We find negative establishment size elasticities unlike the positive estimates for the pooled regression, although the results here are not statistically significant at conventional levels. We still see a positive wage elasticity, although due to the small sample size it is not even significant at the 10 percent level; however, its magnitude is similar to what we see in the pooled regressions in Table 5.

The estimate on the coefficient of bargaining coverage is negative and statistically significant at the 0.9 percent level. The corresponding estimate in Table 4 has a much smaller magnitude and is not statistically significant. However, one might expect the coefficient to be negative. If an establishment has a large fraction of unionized workers (and bargaining coverage were acting as a proxy for unionization), then it could be relatively easy for the establishment to hire new workers as far as recruiting is concerned. A union essentially would do this work for them by choosing the next worker in line for a job.

In Table 7, we repeat the balanced panel regression but now we examine its robustness to outliers, after having examined graphs (not displayed) for observations likely to be exerting undue influence on the estimates. In the first column, we remove only observations whose average log wage changed by at least .5 from 2003 to 2008 in either direction; this represents, approximately, an increase of 65 percent or a decrease of 40 percent in the average wage.[48] It removes only 5 observations from the set of 115 observations that have complete data for the three quantities used, namely log of replacement costs per employee, average log wage and log establishment size for both waves.

Going from Table 6 to the first column of Table 7, the estimates of the log establishment size coefficient stay roughly the same, while estimates of the log wage coefficient drop significantly from 1.84 to 0.18. The second column of Table 7 drops one more observation, namely the only observation with replacement costs per employee changing by more than 5. (It has a value of 9.5, an increase in the value of the replacement costs per employee by a factor of more than 13,000.)[49] In this second column, we see that the estimate of the coefficient on the establishment size changes little, but the estimate for the average log wage is now negative, although far from statistically significant. This casts doubt, however, on our earlier positive wage elasticity estimates.

The final column of Table 7 adds controls to this last regression; it does not exclude any further observations, except those for which there are missing variables. Again, the estimate of the establishment size elasticity is essentially unchanged, and the point estimate of the wage elasticity is negative.

[48] This is not exact because it ignores the fact that the average log wage is averaged over employees within the establishment, but we seek here only to give a sense of how extreme these outliers are.

[49] Here we are removing an observation based on the value of the outcome variable, but it is clearly an unusual observation.

The CES includes data on replacement costs by occupational group and on the entry wage for the most common occupational group at the establishment. This allows us to give an alternate specification, focusing on replacement costs per employee for a particular occupational group. The occupational classifications of professional/managerial workers and blue collar/manual labor employees are the same across waves, although the number of establishments having professionals as the most common category in both waves is quite small. Thus we look only at the blue collar group. We do not have data on recruits by occupation, so we must make a key supposition in order to estimate replacement costs per employee for this group: we assume that recruitment rates are the same establishment-wide as they are within the blue collar occupation.

Table 8 displays regressions of changes between waves in the log of replacement costs for blue collar employees per employee on changes in the log establishment size for this group and on other variables. We use the balanced panel, and restrict to those establishments for which blue collar employees are the most common occupational group in both waves. Column (1) includes only the log of the number of blue collar employees at the workplace as a regressor. The second column adds the log entry pay, while the third adds controls. The fourth through sixth columns are analogous to columns (1)-(3) but with our traditional measure of the log of establishment size replacing the log of the number of blue collar workers.

We find positive point estimates for the size elasticities, for either measure of the number of workers. The magnitudes are smaller than those for the analogous estimates in Table 5. However, all are far from statistically significant at any conventional level, which is not surprising given the small sample sizes. The point estimates for the coefficient on the log of entry pay are negative in columns (2) and (5), which have only size variables and a constant as other regressors; the estimates are not statistically significant however. With added controls, in columns (3) and (6), the estimates are close to zero.

We note finally that similar regressions (not shown here), for each of the blue collar and professional occupation groups, using the pooled sample, give positive size elasticity estimates generally, although they are not statistically significant. The wage elasticity estimates are positive for both groups, and strongly statistically significant for professionals, but not for blue collar employees.

In conclusion, we find evidence from the pooled sample that replacement costs per employee are positively associated with establishment size, and the estimates are strongly statistically significant. Alternate specifications, however, yield point estimates that are close to zero or negative, but none of these are statistically significant at conventional levels.

Regarding the wage elasticity of replacement costs per employee, estimates are positive and strongly significant for the pooled sample. They are also positive and of the same magnitude in a balanced panel regression, albeit not statistically significant. The point estimates are negative when we examine the balanced panel results for robustness to outliers, and finally are either negative or essentially zero for balanced panel regressions among blue collar workers.

6 Interpretation

Now we consider interpretations of the partial correlations we find for replacement costs per employee. We focus here on a key framework for modeling this quantity, due to Manning (2003). We describe the setup and then return to interpreting our results in light of the model.

6.1 Labor Cost Model

In this section, we discuss Manning's labor cost function model. We lay out the intuition in section 6.1.1 and then present a formal model in section 6.1.2.

6.1.1 Motivation

In the simple competitive model and the simple monopsony model, a firm cannot hire any more than $N(w)$ workers at wage w.[50] But by increasing recruitment intensity, a firm may hire (or retain) more workers without increasing w. To embody this notion, Manning defines the *labor cost function,* denoted $C(w, N)$, as the cost *per employee* that a firm must pay, *excluding direct wage costs,* to keep employment at N when paying wage w. We refer to Manning's model as the labor cost function model, after this key component. The idea here is that $C(w, N)$ embodies indirect labor costs, i.e. costs other than the direct wage costs of labor, including replacement costs such as the cost of a worker taking time to become accustomed to a job enough to work at full productivity, the cost of recruiting, selecting and training workers, and other costs such as the cost of firing workers. One would expect that $C(w, N)$ is decreasing in w, as a larger wage would seem to lure more employees and thus might decrease funds spent on recruiting. As we have seen above, empirically there is reason to question this assumption; we discuss this issue in detail below.

Our expectation for the relationship between N and $C(w, N)$ is not as clear; this issue is also a central aspect of our later discussion.

Let us first consider the labor cost functions under the competitive labor market model and under a static monopsony model. In the competitive model, at any wage that is greater than or equal to the competitive wage w_c, a firm can hire as many workers as it wants; however, it can not hire any workers at wages below the competitive wage. Thus, in order

[50] In our discussions of the labor cost function, we draw heavily on Manning (2003) and Manning (2006).

to obtain the standard competitive model as a special case of the larger labor cost model, we set:

$$C(w,N) = \begin{cases} 0 & \text{if} \quad w \geq w_c \\ \infty & \text{if} \quad w < w_c \end{cases} \tag{1}$$

The choice of $C(w,N)$ can be seen in Figure 9. Note that along the line $w = w_c$, the labor cost function jumps.

In order to understand the graph of $C(w,N)$ in the case of monopsony, we recall some key features of the basic monopsony model, which we note makes no (explicit) mention of replacement costs. Figure 11 demonstrates the equilibrium solution in the basic monopsony model. The firm's optimal choice N_0 of N is given by finding the intersection of the marginal revenue product of labor (MRPL) and marginal cost of labor (MCL) curves, and the optimal choice w_0 of w is given by $w_0 = w(N_0)$ on the labor supply curve.

Now consider $C(w,N)$ in the context of the basic monopsony model. At any wage that is above the supply curve, a firm can hire as many workers as it wants; however, it can not hire any workers at wages below the supply curve. Thus, in order to obtain the basic monopsony model as a special case of the larger labor cost model, we set:

$$C(w,N) = \begin{cases} 0 & \text{if} \quad w \geq w(N) \\ \infty & \text{if} \quad w < w(N) \end{cases} \tag{2}$$

This choice of $C(w,N)$ can be seen in Figure 10. Observe that in this graph we have labeled the labor supply curve $w(N)$ in the (w,N)-plane; along this curve, the labor cost function jumps.

6.1.2 Formal Model

Note that in Figures 9 and 10, the labor cost function is clearly not continuous, let alone differentiable. This is presumably not realistic. Moreover, for technical reasons, it is convenient to work with nice functions. Thus, from this point onward, we assume that $C(w,N)$ has continuous second partial derivatives. Thus, maybe the labor cost function in either model might look like a smoothed version of the function in the respective figure. We also make the assumption that

$$C_w(w,N) < 0 \tag{3}$$

holds, which is certainly intuitive, as a larger wage should allow firms to work less hard to attract workers. As mentioned above, once we have described the model in full, we will come back to a discussion of this assumption, in light of our empirical findings. We assume as well that $C_{ww}(w,N) > 0$ holds, which seems plausible, as perhaps there is some decreasing return to lowering the fixed portions of replacement costs that can be gained from raising the wage.

Consider the firm's problem. We assume that the firm has a revenue function $Y(N)$. It costs each firm $w + C(w, N)$ to retain each employee. (Recall that $C(w, N)$ excludes direct wage costs.) Thus the firm's profit function is

$$\pi(w, N) = Y(N) - [w + C(w, N)]N. \tag{4}$$

In this setting, the firm can choose both w and N, as opposed to the competitive or monopsony models in which the choice of one pins the other. The firm chooses w and N to maximize profit. One can think of this as choosing a wage and a recruiting intensity, which then translates into a choice of wage and establishment size. However, consider the function $g(w, N) = w + C(w, N)$. Under suitable conditions, for each fixed value N, there is a unique value of w minimizing $g(w, N)$.[51] Thus it is permissible to define

$$\omega(N) = \min_{w}(w + C(w, N))N, \tag{5}$$

and minimizing $\pi(w, N)$ is then equivalent to minimizing

$$\pi(N) = Y(N) - \omega(N)N. \tag{6}$$

over choices of N. (In a sense, we are first fixing each N, and then choosing w to minimize $(w + C(w, N))$ for that particular choice of N. Then we choose the N that minimizes $Y(N) - [\omega(N)]N$, which is now a function of only N.) Note that the first order conditions are of the form

$$Y'(N) = \omega(N) + \omega'(N)N, \tag{7}$$

which is similar to the first order condition (FOC) in the static monopsony model. We assume that the second order condition holds.

We now note for comparison that the firm's profit in the basic static monopsony model is given by:

$$\pi(N) = Y(N) - w(N)N. \tag{8}$$

Observe that (6) is very similar in form to (8), the expression for profit in the simple static monopsony model, except that here $\omega(N)$ takes the place of the labor supply curve $w(N)$ in

[51] We have $g_w(w, N) = 1 + C_w(w, N)$ and $g_{ww}(w, N) = C_{ww}(w, N) > 0$. Now for any fixed value N_1 of N (excepting implausibly large values of N), we have $\lim_{w\to 0} C(w, N_1) = \infty$ and $\lim_{w\to\infty} C(w, N_1) = 0$; therefore for each fixed N_1, we clearly must have $C_w(w, N_1) < -1$ for some value of w. Thus for each fixed value N_1, there is a solution of $g_w(w, N_1) = 0$. As, for each fixed value N_1, we have that $g_{ww}(w, N_1) > 0$ for all choices of w, $g(w, N_1)$ is globally convex as a function of w, and therefore any solution of $g_w(w, N_1)$ globally minimizes $g(w, N_1)$. (Here N_1 is still fixed; we are not implying that this solution minimizes over all choices of w and N, just over all choices of w with the fixed choice N_1.)

(8). So a key issue is whether or not $\omega(N)$ is increasing in N. If $\omega(N)$ is increasing in N, the equilibrium analysis looks much like that of the static monopsony model in Figure 11, with $\omega(N)$ taking the place of $w(N)$. If $\omega(N)$ is, say, constant in N, then the equilibrium analysis is akin to that in the competitive model. Thus we are interested in whether or not $\omega(N)$ is increasing.

To investigate this further, we use the Envelope Theorem to note that

$$\omega'(N) = C_N(w, N). \tag{9}$$

Whether this model looks like the monopsony model or the competitive model hinges on whether or not $C(w, N)$ increases with N.

6.2 Interpretation of Results in Context of the Manning Labor Cost Model

The labor cost function should be a measure of the indirect labor costs per employee, i.e. the establishment's total labor costs other than the direct wage costs of labor, divided by the number of workers. We can view our results in terms of the Manning model, by taking the quantity of replacement costs per employee as our operational definition of the indirect labor cost per employee.

Under the model, whether or not the results are consistent with monopsony depends on whether or not the function $C(w, N)$ is increasing in the establishment size N or constant in N. One way of implementing this in practice of course is to regress $C(w, N)$ on w and N and test whether or not the coefficient on N is positive. Obviously endogeneity issues may arise.

But we have performed these regressions in section 5.3. There we saw evidence from pooled regressions that replacement costs per employee are positively correlated with establishment size, i.e. that $C(w, N)$ is increasing in N. Moreover, these results are statistically significant at the 0.4 percent level or lower. This provides some evidence for the monopsony model. On the other hand, alternative specifications yield point estimates that are nearly zero or are negative, although these are not statistically significant at any usual levels. Thus the evidence is mixed.

In comparison, Manning (2006) attempts to estimate the labor cost function using United Kingdom Institute for Personnel Development data on turnover rates.[52] The data include information on turnover costs for ten occupational groups, as well as turnover rates. Wages are not given, however, and there is little information on characteristics of the establishment. Manning regresses the log of turnover costs per employee on logs for the regressor variables.

[52] This is also discussed on pages 292-296 of Manning (2003).

Manning assumes that the turnover rate can be written as a linear function of the log wage. Using this and the FOC (7), one can then remove both the wage and the number of recruits from an expression for the log labor cost function. Manning then is able to estimate a regression of log labor costs on just the turnover rate and log employer size. The procedure does not allow for estimation of the original coefficient on employer size, but can test whether or not it is positive. He finds positive estimates and is able to reject the null hypothesis that the estimates are zero.

Returning to our work, consider one of the key assumptions of the labor cost model, namely (3), which posits that the labor cost function $C(w, N)$ is decreasing in the wage w. Our estimates in a pooled regression of the log of replacement costs per employee on the average log wage were actually *positive,* and statistically significant at the 0.2 percent level or lower, even in a specification with multiple covariates and industry indicators as controls. The wage elasticity estimates were also positive, and of at least the same magnitude, in panel regressions, although not statistically significant at standard levels, perhaps because of the small sample size in the balanced panel. We also reported two specifications in which the wage elasticity estimates were negative or essentially zero.

Our results suggest that the wage elasticity may in fact be positive. This presents a puzzle: how can one reconcile this (possible) positive relationship with the intuition that an employer should be able to reduce recruiting effort if the establishment increase wages? We offer some initial thoughts in the next section.

6.3 Discussion of Mixed Results on Wage Elasticity

We consider a few observations that may explain why our estimates are not consistent with the labor cost model's wage elasticity assumption.

6.3.1 Heterogeneity

As noted above, our findings of a partial correlation between the labor cost function $C(w, N)$ and the wage w do not imply that the first derivative of $C(w, N)$ with respect to w is positive, as the estimates may be biased, and the estimators may not even be consistent. We consider a few reasons why this may be the case.

One issue is unobserved heterogeneity. It could give rise in the data set to a spurious positive association between the wage and labor costs. Consider the well-known setup which demonstrates this type of heterogeneity problem at a basic level: one could have groupings of establishments within which the labor cost function is indeed decreasing in the wage, but when putting these groupings together, one sees a positive association between wage and the labor cost. Here the picture of the observations in a plane is of clusters of points such that there is a clear downward trend within each cluster, but the clusters are themselves arranged along an increasing line, thus leading to a positive association when one calculates regression estimates. There could be more complicated versions of the same setup.

Our first method for confronting the issue of heterogeneity was to control for multiple characteristics of an establishment, including industry fixed effects, in the pooled regressions. Table 6 addressed heterogeneity by using the balanced panel to control for unobserved establishment characteristics.[53] We found similar wage elasticity estimates, although not statistically significant ones; but the sample sizes were quite small. Table 7 showed that these results are not robust to outliers.

In Table 8, we focused on a more homogeneous set of workers, by restricting the sample to establishments with predominantly blue collar workers. We found wage elasticity estimates close to zero or negative, and none were statistically significant. Although the basic model assumes for clarity that all wages are the same, the entry wage is the particular wage that one would expect to decrease labor costs through an increase. Since this set of regressions uses *entry* wages, they are closer in some sense to the spirit of the labor cost function model.

In summary, despite our attempts to control for heterogeneity, we do find some evidence for a negative wage elasticity of labor costs, but also somewhat stronger evidence for a positive elasticity.

6.3.2 Measurement Error and Other Sampling Bias

An alternate explanation for finding mixed results on the wage elasticity involves the relationship of the survey question to the theoretical notion of indirect labor costs. When the survey asks about replacement costs, it offers "on-the-job learning" as one of the specific examples of such a cost. Recall equation (4): in the labor cost model, the revenue function does not depend on the skill level of the employee. Thus, within the framework of the model, the loss to an establishment from employing an employee who is not yet fully productive should be included in indirect labor costs. But the wage cost of training should not be included in indirect labor costs, as this is captured in the direct wage cost term wN. It is only the further loss to the establishment from the reduction in production due to lower productivity that should be included.

It is unclear whether a survey participant would include this further loss in an estimate of replacement costs or, on the other hand, would simply make an estimate of the full cost of training and include this cost in a response for replacement costs. The latter would lead to an overestimate of indirect labor costs per employee, as it ignores that workers could be producing at least some good or service during training; for our purposes, they are in effect

[53]Although this is a strong design, it need not necessarily capture all unobserved heterogeneity, as noted above.

double-counting.

To investigate this issue, we first estimate the production value of the workers during the training period for new employees. Assume for the sake of this estimation that the marginal revenue product of labor equals the wage. If employees produce during training at a portion of full productivity that starts at zero and linearly reaches full productivity at the end of the training period, then the new employees' production value during the training is one half of the product of the wage and the number of new training hours reported by respondents.[54] This is the amount by which survey participants would be overestimating indirect labor costs per recruit in their responses, under our above assumptions. We subtract the result of this calculation from the replacement costs reported, take the maximum of this difference and zero, and finally multiply by the (5-year) ratio of recruits to employees. We call this value the replacement costs net of production value per employee.

Table 9 reports results from regressing the log of this quantity on log establishment size and the log wage.[55] As above, the first column uses only the log establishment size and a constant as regressors, while the second column adds the log wage. The final column includes multiple control variables.

The point estimates of the coefficient on the log wage are negative. They are not statistically significant at any usual levels, but the $p-$value for the final specification is only 0.2, which is not that large, given the small sample size. Thus it appears that one could find negative wage elasticity estimates with a panel sample in which one asked a question designed specifically to estimate indirect labor costs, as opposed to the question in the CES that we use, which was not explicitly designed for this purpose. We emphasize that the outcome variable in the regression is imputed under strong assumptions, and we use it only to give a sense of whether or not one might obtain a negative wage elasticity estimate.[56] We note also that a similar regression, but with the pooled sample, which we do not report, still gives positive wage elasticity estimates. So it need not be the case that a different question would yield negative estimates, but as seen from the above, it could give that result.

A further type of measurement error could occur if many respondents give a very quick estimate of replacement costs, in which they simply multiply the wage by a common estimate of the number of training hours required. This form of measurement error would lead to a strong mechanical pull, among these observations at least, towards a positive wage elasticity estimate.

Nonresponse bias could also lead to our positive wage elasticity results. Nonresponse can

[54]See the data appendix for more detail for a similar argument.

[55]As above, we use the log of one plus the quantity rather than the log.

[56]Of course, one expects the wage elasticity to be lower than the values in Table 6 because of the mechanical effect of subtracting a product of the wage from the outcome variable: our purpose here is to present an idea of the magnitude of the effect that could occur.

occur at the observation level, and also at the level of the variable, for an establishment that does respond to at least some of the questions. It is not clear which way this bias would lead.

6.3.3 Decomposition of Wage Elasticity Estimates

Tables 10 and 11 investigate further the finding of a positive wage elasticity in the pooled sample. Here we decompose the wage elasticity estimate in the original pooled regressions in Table 5 into two parts: the wage elasticities of each of the log replacement costs per recruit and the log of the recruitment rate (i.e. the log of the ratio of recruits to employees).

We compare the results in these two tables with the results in Table 5. In these two tables, we have restricted to observations for which all variables required for Table 5 are not missing so that the sample does not change, in order to yield better comparisons.[57]

For technical reasons related to normalizing the log variables, the coefficients in Tables 10 and 11 do not sum to the corresponding coefficient in Table 5, as one might expect from the linearity of the ordinary least squares estimator in the outcome variable. For these reasons, the magnitudes of the estimates in the tables are not directly comparable.[58] Nevertheless, we can still get a sense of the sign of these coefficients.

In both Tables 10 and 11, as in Table 5, the first column reports a regression on only the log establishment size and a constant, while the second column adds the average log wage, the third column adds various controls and the final column includes 1-digit industry controls and drops the constant term.

We see from Table 10 that the log of the replacement cost per recruit is positively associated with the establishment size and the wage. On the other hand, Table 11 shows that the log of the recruitment rate is negatively associated with both the log establishment size and the log wage (although weakly so with the wage). Thus there are positive associations

[57] Because the variables in Tables 10 and 11 are either used in Table 5 or are intermediate variables required in the construction of the log of replacement costs per employee, the lack of missing variables in Table 5 for a given observation implies there are no missing variables for that particular observation in the other two tables.

[58] The issue is that the logarithm of the labor cost is not precisely equal to the sum of the log of each of the component factors, because in certain cases, we must look at the log of the sum of a constant and the particular factor, so that we do not have an undefined variable when the factor is zero. This actually causes the sum to be significantly different from the coefficient in Table 5, especially because the log of the recruitment rate is much smaller in magnitude because we have chosen to add a one to it before taking the log. If we had chosen a number such as .001 to add to the raw recruitment rate before taking logs, the logs would be negative and much larger in magnitude, in general, and thus we would expect to see larger changes in the log of the recruitment rate associated with changes in the log of the wage. We could have added a number such as .001 instead in order to make the coefficients sum more nicely, but this would come at the price of also making the standardization of the log of replacement costs per employee much more complicated, as, for example, it would involve a term that is a product of .001 and the replacement costs per employee. We have chosen to make the modification in the log of replacement costs per employee more standard over the goal of clean comparison in these three tables.

with wage and size coming from the replacement costs per recruit, and negative associations from the recruitment rate. We can see from Table 5 that the former association dominates the latter. The fact that the recruitment rate is negatively associated with the wage implies, at equilibrium, that the quit rate decreases as the wage increases.[59] Similarly the quit rate is decreasing in the establishment size. At the same time, replacement costs per recruiting event are increasing in both the wage and the establishment size.

Replacement costs per recruit are larger when establishments offer health insurance, perhaps because of the administrative costs of setting up health insurance plans, as suggested above. The costs per recruit are negatively associated with the fraction working part-time, which one would expect if employers invest less in training such workers, given that any benefits of training will be lower for those working fewer hours per week. However, some of the other estimates have a surprising sign or are unexpectedly not statistically significant at usual levels. The recruitment rate is negatively associated with bargaining coverage, and positively associated with the fraction of part-time workers. These results are to be expected if one makes the assumption that hiring takes place in an equilibrium framework, so that the recruitment rate equals the quit rate. But under this assumption, it is surprising that having a higher fraction of college graduates is not negatively associated with recruitment rates and offering health insurance is not strongly negatively associated with the recruitment rate.

6.3.4 Model Incompleteness

Consider another possible factor that could be playing a part in the positive wage elasticity findings. In the profit function (4), the revenue function is assumed not to depend on the wage. It could be the case in practice that a decrease in wages may lead to a lower value of the labor cost function, but would reduce production at the same time. If so, the model would not accurately capture this aspect of the workplace.

This situation could occur if a lower wage reduces effort levels, as discussed in the work of Bewley (1999), and in efficiency wage models, e.g. the model of Shapiro and Stiglitz (1984). Here we are not thinking of a situation in which workers completely shirk as in the Shapiro-Stiglitz model, and thus may get fired, which would affect the recruitment rate, and thus the labor cost function. Rather, in this situation, workers only reduce effort, which affects the revenue function.

The exposition in section 6.1.2 reveals that the assumption that replacement costs are decreasing in the wage is important for the model. If these costs were increasing in the wage, then the optimal choice of the wage would be zero. But if a model were to include an effect of the wage on the revenue function, then this model could be consistent with a

[59] Naturally, this is not necessarily as a causal relationship.

positive association between replacement costs and the wage. This reasoning illustrates why one might expect to find a positive wage elasticity if a model were to capture more aspects of the employment decision than the labor cost model.

7 Conclusion

In this paper, we study employee replacement costs and their relationships with multiple workplace characteristics. In the California Establishment Survey (CES), we have an excellent new data set to study such questions. It contains data on replacement costs, establishment size, wages, and numerous relevant establishment characteristics. Moreover, it has a two-wave panel structure, which allows us to control for unobserved qualities of each workplace, although the size of the balanced panel sample limits the statistical power.

We note the substantial size of replacement costs per recruiting instance, and the large variation occurring in this quantity. We also find that replacement costs are negatively associated with establishment policies of prioritizing seniority when filling a position from within. We posit that this association could arise from savings on training costs from hiring more senior workers, or more likely from lower hiring transactions costs from using seniority as a simple promotion criterion.

We also investigate replacement costs per employee, a measure that combines the replacement costs per recruit and the employee turnover rate. We find some evidence for positive partial correlations between replacement costs per employee and both the establishment size and the wage; however, the relationship is robust in neither case, but in both cases the evidence is somewhat stronger for a positive correlation.

Each of these relationships plays an important role in a model due to Manning that incorporates replacement costs. The central idea of this labor cost model is that in addition to varying the wage in order to hire more workers, firms can increase some types of recruiting effort without changing the wage, and in this fashion recruit more employees. A key assumption of the labor cost model is that an increase in the wage should allow an employer to reduce recruiting costs.

This runs counter to our finding that the replacement costs per employee may in fact be *positively* correlated with the wage. We have offered some ideas about how one might reconcile this finding with the model, but we cannot conclusively explain it. It stands as a challenge for future researchers.

A Data Appendix

A.1 Wage Measures

We define the various wage measures used. Two measures used are the average wage or average logarithm of the wage at the establishment. These are calculated using responses to questions about the fraction of workers in certain wage categories and assigning the midpoints of the wage categories (or, respectively, the average of the log at each of the endpoints) to that fraction of workers. In the top wage bracket, the average wage in that category is estimated assuming a Pareto distribution. (The log wage in that category is taken to be the log of that average.) Another wage measure used is the hourly entry wage in the establishment's most common occupational category.

All dollar measures are in 2003 dollars; 2008 measures are adjusted using the annual levels of the Consumer Price Index for all Urban Consumers (CPI-U) for 2003 and 2008. Also, we actually more often use logarithms than levels; we note that in cases where the quantity is often 0, we actually use the logarithm of the sum of one and the relevant quantity, yet still refer to it as just the logarithm of the quantity.

A.2 Imputed Replacement Costs

We describe how we estimate what we call *imputed replacement costs*. By imputed replacement costs, we mean that we use the respondents' answers if nonmissing and nonzero; if on the other hand, the answers are missing or zero, we use an imputed value of replacement costs in their place, assuming the imputed measure itself is nonmissing. We use the following procedure to estimate these imputed replacement costs. We describe the procedure for the 2003 wave and then note the mild variants used for the 2008 wave.

First we impute a variable that is the number of working hours to full productivity for the average employee. This is asked, in months, for professional employees in one question, and for blue collar and clerical workers together in another question. We assume that sales workers have the same time to productivity as blue collar and clerical workers. We combine to get an average over all workers using respondents' answers to questions about the fraction of workers in each occupational groups. (The latter answers are normalized so that they sum to 1.) We also assume that there are 160 working hours in each month.

Then, given the number of working hours to full productivity, we make the strong assumption, for the sake of making an estimate, that new workers' productivity rises linearly from 0 to full productivity, over the course of the time to full productivity. Moreover, we assume that workers are paid their marginal product and thus that the establishments lose in wages the difference between full productivity and current productivity.

Assume that the wage is constant over the time to full productivity. As the wages paid are a rectangle with height equal to wage and length equal to the hours to full productivity, the loss to the establishment is the area of a triangle that is half the area of the rectangle.

Concerning the new training hours, we assume that the entirety of the wages for new workers for the new training hours are lost to the establishment. There is obviously some double-counting here, vis-a-vis the losses due to lack of full productivity, but here any losses due to training by managers or other workers counteract that. One can see that, out of necessity, this is a very rough approximation.

We add up the losses due to lower productivity and new training hours to get an imputed replacement cost for each establishment. We note that the procedure differs slightly for 2008 in a few ways. For one, questions on the time to full productivity were asked in a different way. Rather than being asked a few possibilities for how many months it took, respondents were asked if it took less than one month, and if not less, were then asked to state the number of months. We note that this seems to have led to much different answers. Also, in 2008, respondents were asked a question on the time to full productivity for nonprofessional and nonmanagerial employees, as opposed to, in 2003, a question for blue collar and clerical workers. Finally, as with other dollar figures, we adjust for inflation, so all figures are in 2003 dollars.

We can compare these imputed replacement costs with our preferred replacement cost measure that is calculated more directly from survey responses. From Table 2, we see that their means, and the means of their logs, are fairly close to each other. Correlations between the two measures, among all observations for which both are defined, also give a sense of the relationship between them. In the pooled sample, the correlation is .23, which is perhaps lower than one might expect. Breaking down by wave, the correlation is lower in 2003, at .17, but much higher in 2008, at .44.[60] Thus there are indications, from the comparison of means and from the 2008 correlation, that our imputation process captures a good portion of actual replacement costs, but overall, the evidence leans against this conclusion.

[60] These correlations use the survey weights. The respective sample sizes are 808, 258 and 550.

References

[1] **Abowd, John and Francis Kramarz,** "The Costs of Hiring and Separations," *Labour Economics,* Oct. 2003, 10, 499-530.

[2] **Anderson, Patricia,** "Linear Adjustment Costs and Seasonal Labor Demand: Evidence from Retail Trade Firms," *The Quarterly Journal of Economics,* Nov. 1993, 108, 1015-1042.

[3] **Bewley, Truman F.,***Why Wages Don't Fall During a Recession,* 1999, Harvard University Press, Cambridge, MA.

[4] **Blatter, Marc, Samuel Mühlemann and Samuel Schenker,** "The Costs of Hiring Skilled Workers", Nov. 2009, Swiss Leading House Working Paper No. 15, available from http://www.isu.unizh.ch/leadinghouse/WorkingPapers/0015_lhwpaper.pdf , downloaded on January 31, 2010.

[5] **Campbell, Carl M.,** "Do Firms Pay Efficiency Wages? Evidence with Data at the Firm Level," *Journal of Labor Economics,* 1993, 11, 442-470.

[6] **Cascio, Wayne F.,** *Managing Human Resources: Productivity, Quality of Work Life, Profits,* 7th ed. 2006, Mc-Graw Hill/Irwin, New York.

[7] **Fairris, David,** "Internal Labor Markets and Worker Quits," *Industrial Relations,* 2004, 43, 573-594.

[8] **Hamermesh, Daniel,** *Labor Demand*, 1993, Princeton University Press, Princeton, New Jersey.

[9] **Hamermesh, Daniel and Gerard Pfann,** "Adjustment Costs in Factor Demand," 1996, *Journal of Economic Literature,* 34, No. 3, pp. 1264-1292.

[10] **Kramarz, Francis and Marie-Laure Michaud,** "The Shape of Hiring and Separation Costs in France," *Labour Economics,* 2010, 17, 27-37.

[11] **Manning, Alan**, *Monopsony in Motion,* 2003, Princeton University Press, Princeton, New Jersey.

[12] **Manning, Alan**, "A Generalised Model of Monopsony", 2006, *The Economic Journal,* 116 (January), 84-100.

[13] **Mas-Colell, Andreu, Michael Whinston and Jerry Green**, *Microeconomic Theory*, 1995, Oxford University Press, New York.

[14] **Mortensen, Dale and Éva Nagypál,** "More on Unemployment and Vacancy Fuctuations," 2007, *Review of Economic Dynamics*, 10, 327347.

[15] **Mortensen, Dale and Christopher Pissarides,** "New Developments in Models of Search in the Labor Market," in O. Ashenfelter and D. Card (eds.), Handbook of Labor Economics 3, 1999, 2567-2627, Amsterdam: Elsevier.

[16] **Oi, Walter Y.,** "Labour as a Quasi-Fixed Factor," *Journal of Political Economy,* 1962, 70, 538-55.

[17] **Osterman, Paul,** "How Common is Workplace Transformation and Who Adopts it?" *Industrial and Labor Relations Review,* January 1994, 47, 173-188.

[18] **Paradise, Andrew**, "Learning Remains Steady During the Downturn", *Training and Development,* November 2009, American Society for Training and Development, from http://www.astd.org/TD/Archives/2009/Nov/Free/0911_SOIR.htm and http://www1.astd.org/TDimages/2009/nov/SOIR-Chart3L.png , accessed January 31, 2010.

[19] **R Development Core Team**, *R: A Language and Environment for Statistical Computing,* 2009, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org .

[20] **Shapiro, Carl, and Joseph E. Stiglitz,** "Equilibrium Unemployment as a Worker Discipline Device," *The American Economic Review,* 1984, 74, 433-444.

[21] **Silva, José Ignacio and Manuel Toledo,** "Labor Turnover Costs and the Cyclical Behavior of Vacancies and Unemployment," 2009, *Macroeconomic Dynamics*, 13 (Supplement 1), 76-96.

[22] **Solomon, Jolie,** "Companies Try Measuring Cost Savings From New Types of Corporate Benefits", *The Wall Street Journal,* December 29, 1988, p. B1.

[23] **Survey Research Center, University of California, Berkeley,** *2003 California Establishment Survey,* 2003, sponsored by the Institute for Research on Labor and Employment, University of California, Michael Reich, Principal Investigator.

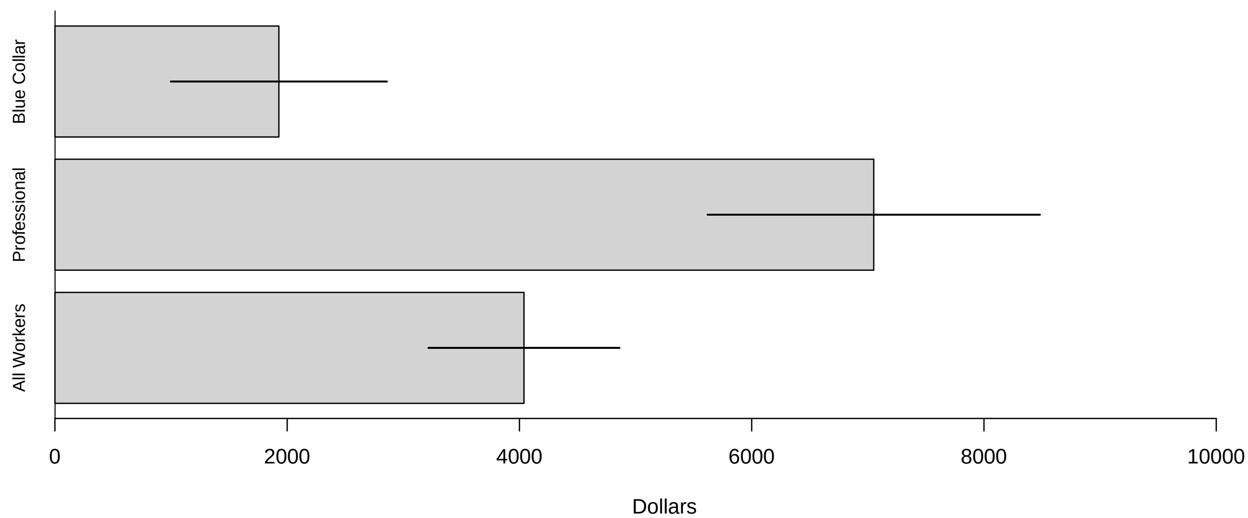
[24] **Survey Research Center, University of California, Berkeley,** *Codebook for 2003 California Establishment Survey,* 2003.

[25] **Survey Research Center, University of California, Berkeley,** *2008 California Establishment Survey,* 2008, sponsored by the Institute for Research on Labor and

Employment, University of California, Michael Reich and Sylvia Allegretto, Principal Investigators.

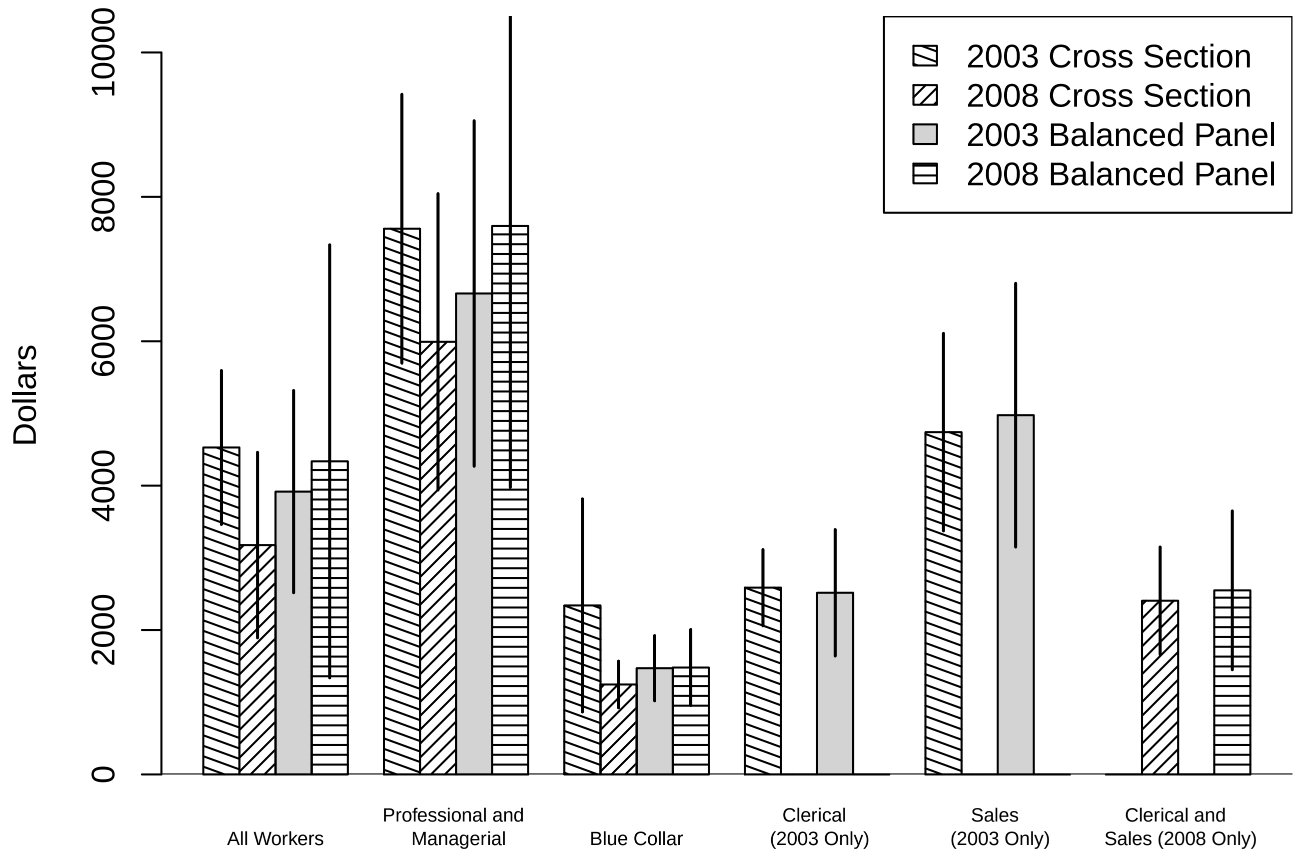
[26] **Survey Research Center, University of California, Berkeley,** *Codebook for 2008 California Establishment Survey,* 2009.

Figure 1: Mean Replacement Costs (per Recruit)



Notes: Calculations use pooled sample of 2003 and 2008 waves of the CES, and are clustered at the establishment−level. Costs are in 2003 dollars. Length of lines at right side of bars represent 95% confidence intervals for the mean. Calculations use establishment weights adjusted for nonresponse and calculations of standard errors account for the stratified nature of the sample. Note that none of the replacement costs are strictly comparable across years, as the categories about which replacement cost questions were asked changed across waves. However, in both waves, questions were asked about replacement costs for professional/managerial workers as well as blue collar/manual labor workers; on the other hand, in 2008, clerical and sales were grouped together.

Figure 2: Mean Replacement Costs, by Survey Wave



Notes: Costs are in 2003 dollars. Lines at the top of each bar represent 95% confidence intervals for the mean. Note that none of the replacement costs are strictly comparable across years, as the categories about which replacement cost questions were asked changed across waves. However, in both waves, questions were asked about replacement costs for professional/managerial workers as well as blue collar/manual labor workers; on the other hand, in 2008, clerical and sales were grouped together. 2003 Cross Section and 2008 Cross Section refer to the full data sets for those years, while the 2003 Balanced Panel and 2008 Balanced Panel refer to the subset of establishments present in both waves and asked the relevant question in both years. Calculations for the 2003 and 2008 cross sections use establishment weights adjusted for nonresponse for the respective year, while calculations for the 2003 and 2008 balanced panel subsets use establishment weights adjusted for nonresponse for the 2003 wave and adjusted further for nonresponse for the 2008 wave. For both the 2003 and 2008 balanced panels, 2003 strata are used when accounting for stratification.

Figure 3: Histograms of Replacement Costs (See text for details.)

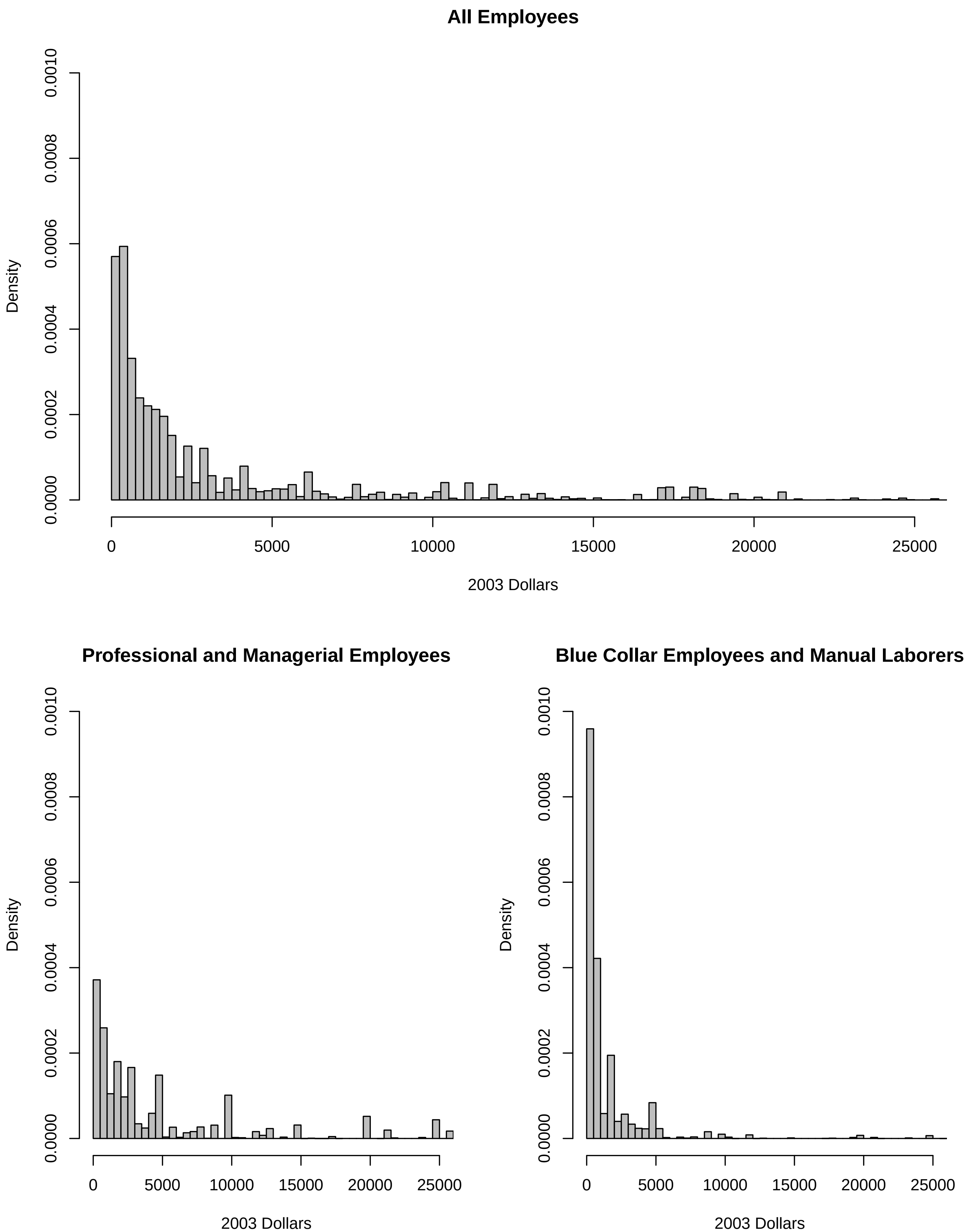
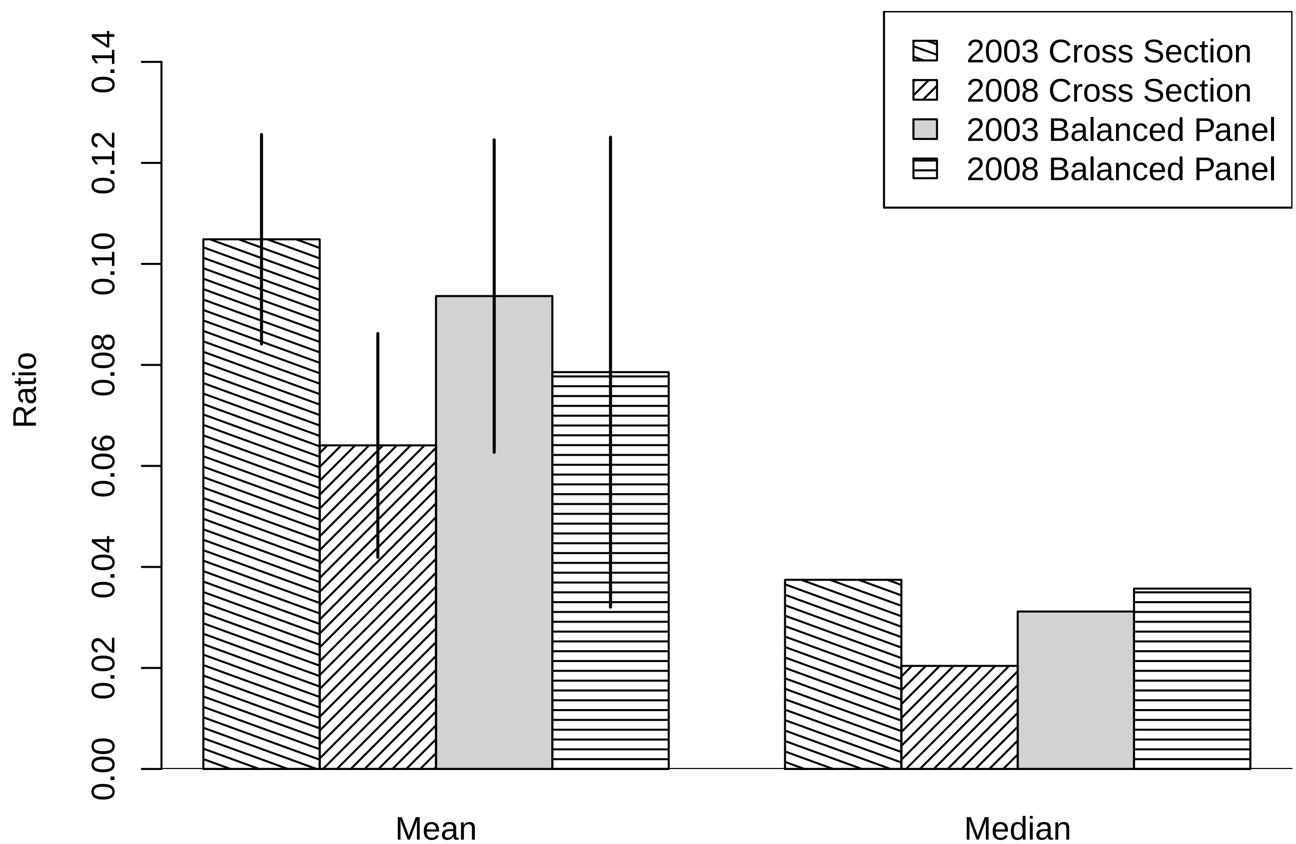


Figure 4: Mean and Median of Ratio of Replacement Costs to Annual Wage



Notes: Lines at top of bars represent 95% confidence intervals on either side of the mean. Workers are assumed to work 2000 hours per year. Average wage estimated using midpoints of response ranges. Note that none of the replacement costs are strictly comparable across years, as the categories about which replacement cost questions were asked changed across waves. However, in both waves, questions were asked about replacement costs for professional/managerial workers as well as blue collar/manual labor workers; on the other hand, in 2008, clerical and sales were grouped together. 2003 Cross Section and 2008 Cross Section refer to the full data sets for those years, while the 2003 Balanced Panel and 2008 Balanced Panel refer to the subset of establishments present in both waves and asked the relevant question in both years. Calculations for the 2003 and 2008 cross sections use establishment weights adjusted for nonresponse for the respective year, while calculations for the 2003 and 2008 balanced panel subsets use establishment weights adjusted for nonresponse for the 2003 wave and adjusted further for nonresponse for the 2008 wave. For both the 2003 and 2008 balanced panels, 2003 strata are used when accounting for stratification.

Figure 5: Histogram of Training Hours for New Employees, Pooled Sample

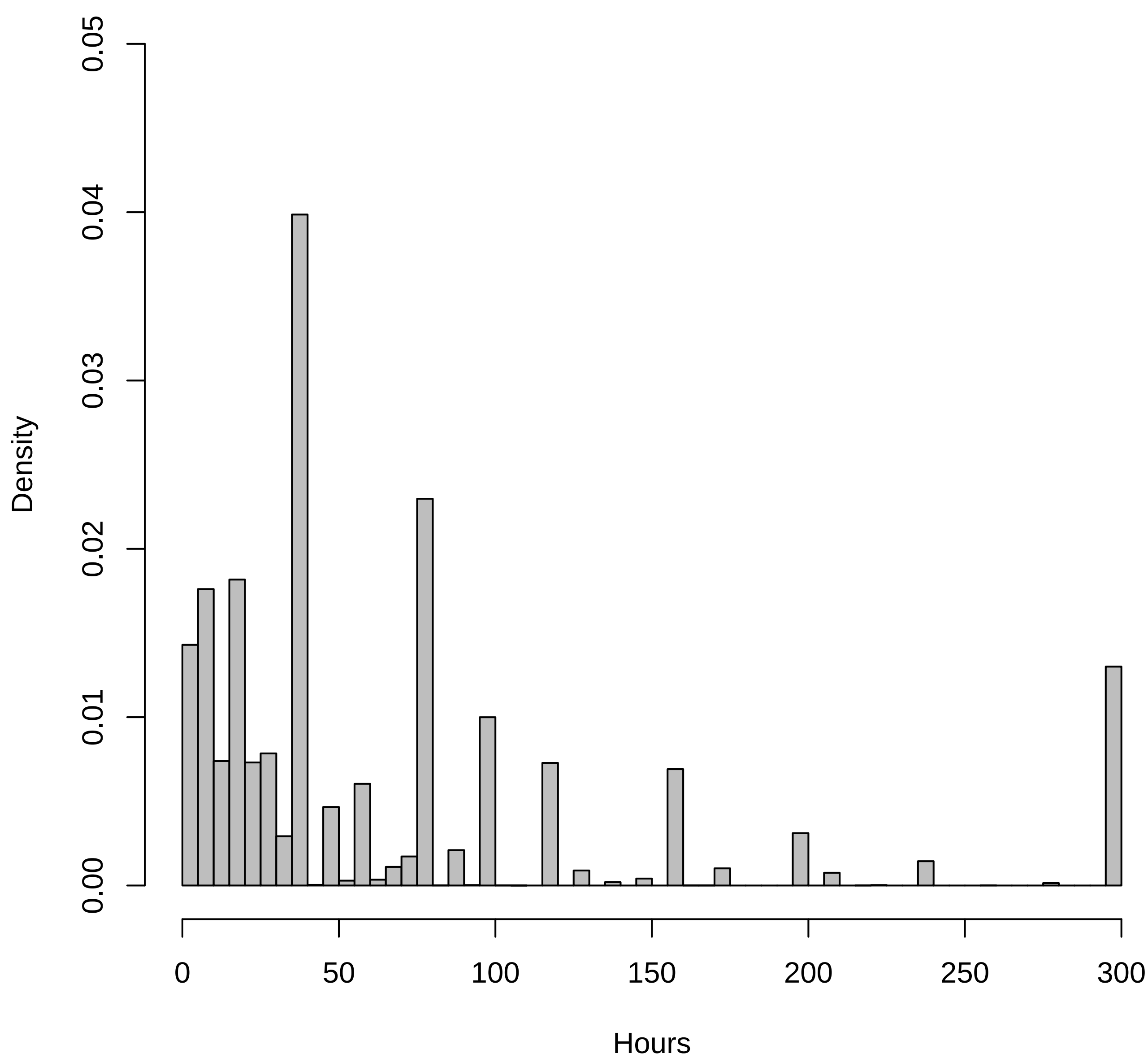
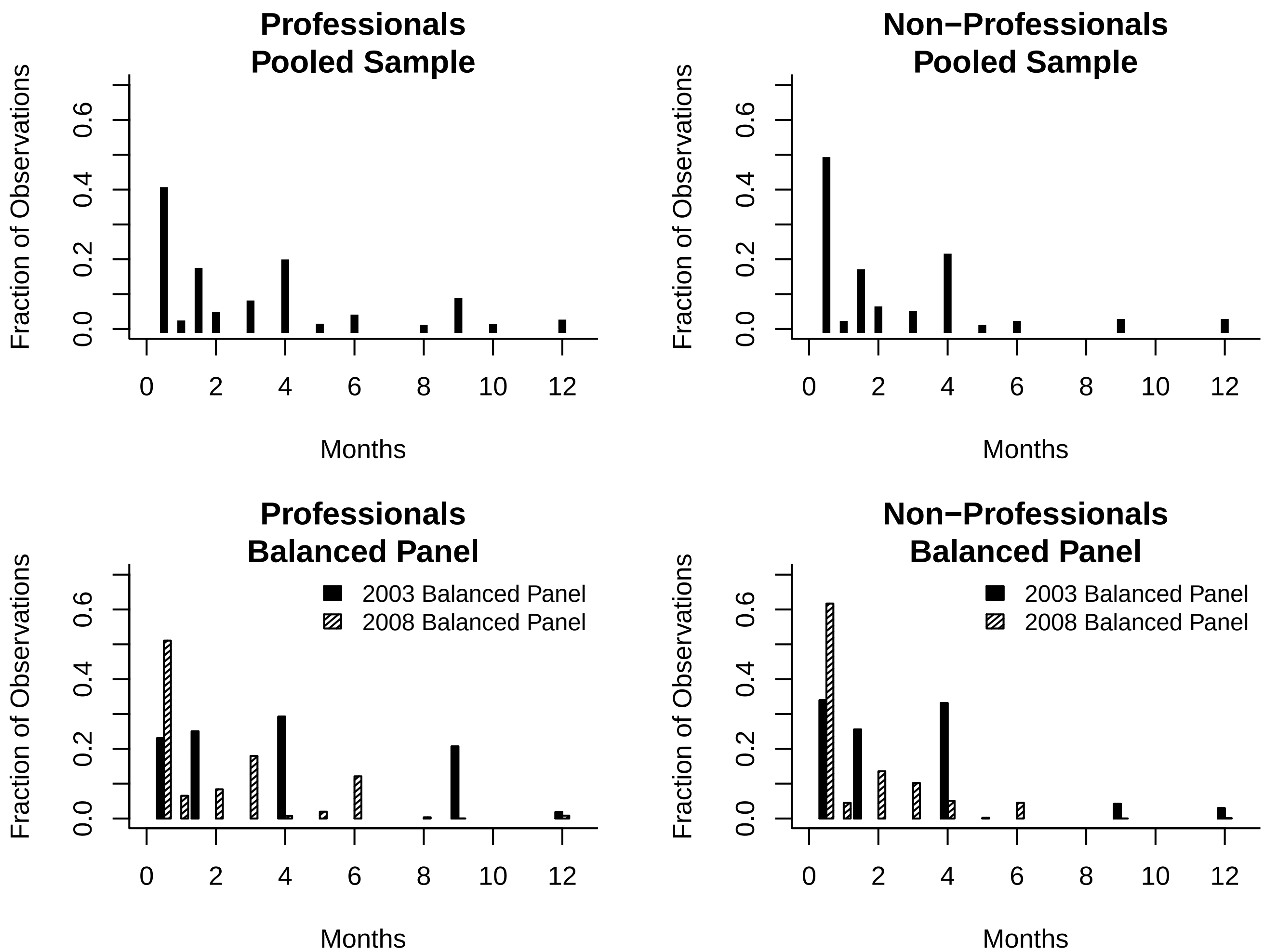


Figure 6: Histograms of Time to Full Productivity for New Employees



Note that the graphs are cut off at 12 months to improve readability, which excludes some responses. Note that the questions were asked in different ways across waves. Number of months represent midpoints of categories in 2003: 0.5 represents less than a month; 1.5 for 1−2; 4 for 2−6; 9 for 6−12; and 12 for a year or more. In 2008, on the other hand, 0.5 represents less than a month, but 1 or greater represents that number of months. The non−Professional category is not strictly comparable across waves.

Figure 7: Scatter Plot of Log of Replacement Costs per Employee by Log of Establishment Size

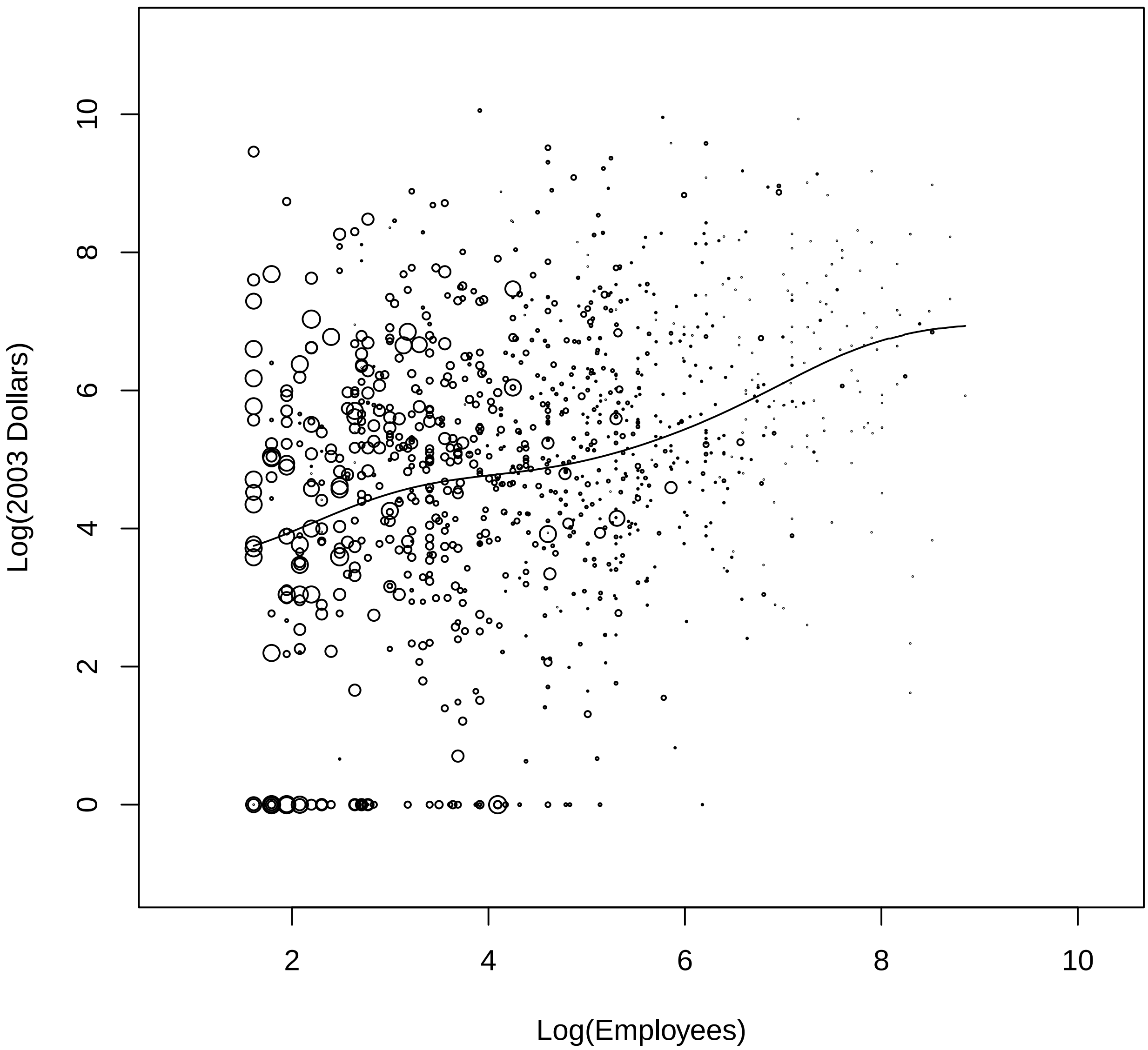


Figure 8: Scatter Plot of Log of Replacement Costs per Employee by Log Wage

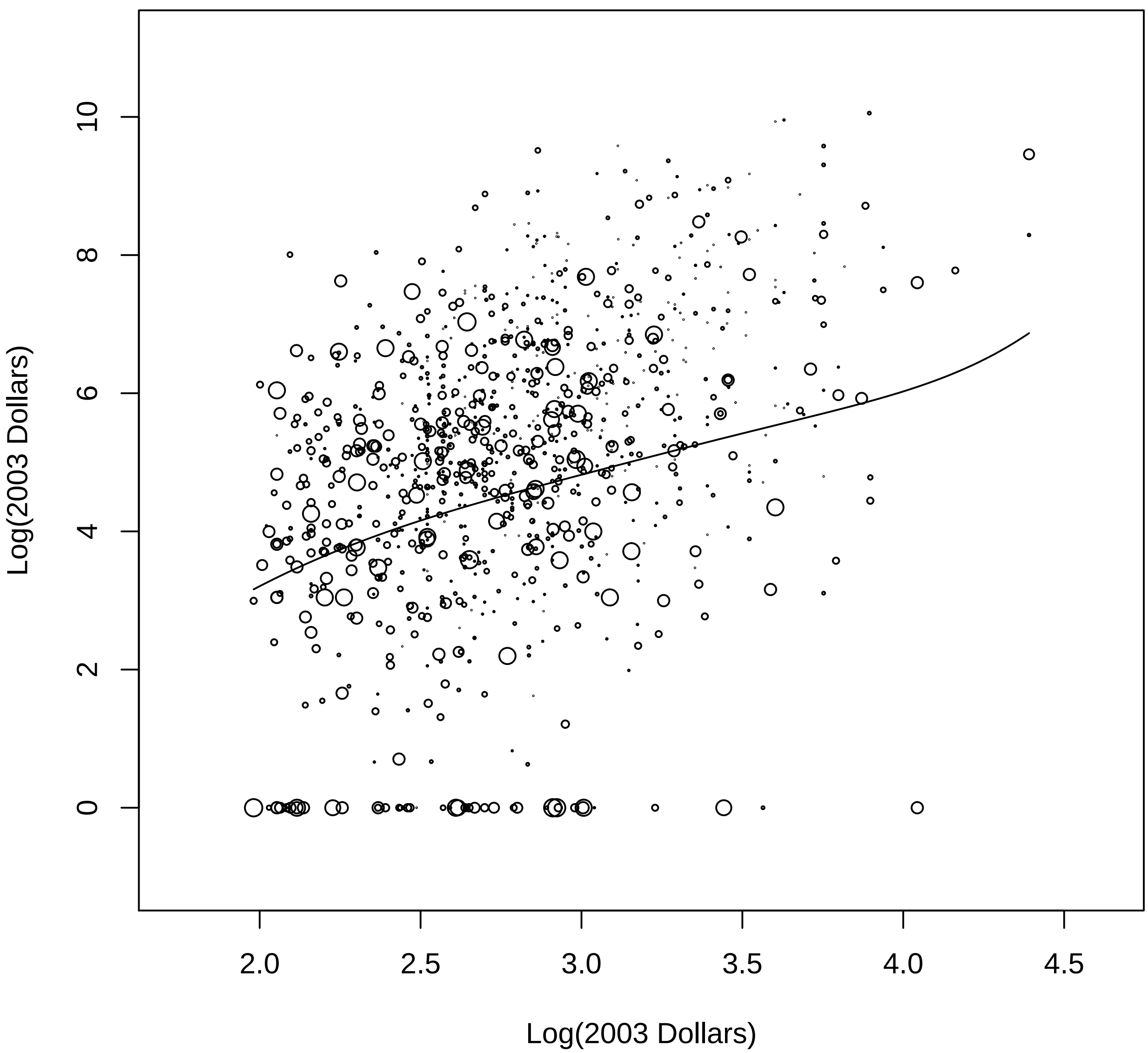


Figure 9: The Labor Cost Function for the Basic Competitive Model

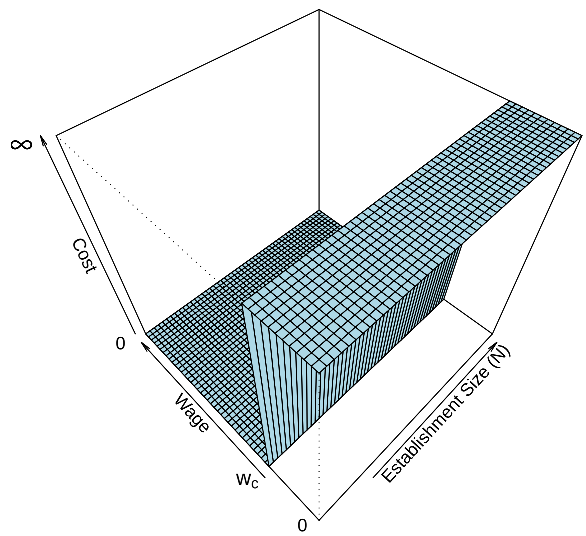


Figure 10: The Labor Cost Function for the Basic Monopsony Model

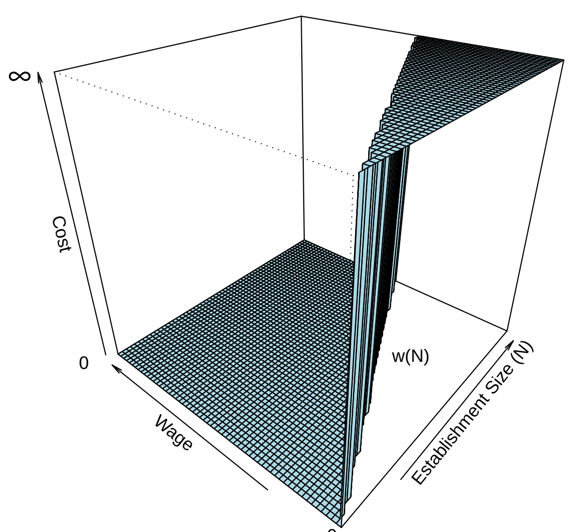


Figure 11: The Static Monopsony Model

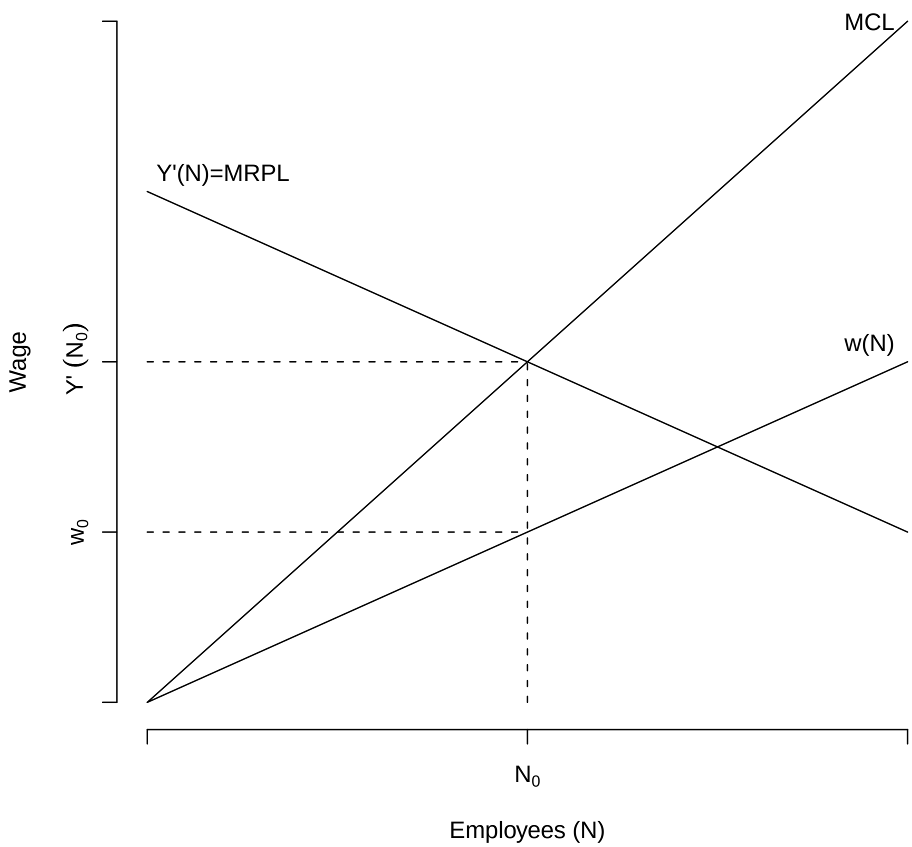


Table 1: Sampling Results and Response Rates in the CES

	2003	2008 Panel Sample	2008 New Sample
Sampled Establishments	2806	1080	1072
2008 Panel Sample Still in Business in CA		1016	
Number Meeting Eligibility Criteria	2200	868	849
Interview Completed	1080	358	294
Response Rate (Unweighted)	49.1%	41.2%	34.6%

Table 2: Descriptive Statistics for Key Variables, Pooled Sample

	Mean of Level (Std. Error)[Std. Dev.]	Mean of Log (Std. Error)[Std. Dev.]
Replacement Costs Per Recruit	4039.01 (419.7)[9799.8]	6.41 (0.2)[2.7]
Establishment Size	43.83 (2.81)[175.73]	2.87 (0.04)[1.07]
Recruits (1 Yr Rate)	6.9 (0.41)[29.79]	1.21 (0.04)[1.1]
Recruits (5 Yr Rate)	4.38 (0.26)[17.33]	1.11 (0.03)[0.81]
Average Hourly Wage	18.55 (0.52)[10.26]	2.67 (0.02)[0.43]
Hourly Entry Wage	13.19 (0.46)[9.13]	2.43 (0.03)[0.51]
New Training Hours	65.59 (4.1)[76.25]	3.48 (0.08)[1.42]
Fraction Professional or Managerial	0.28 (0.01)[0.25]	
Fraction Blue Collar	0.37 (0.02)[0.33]	
Fraction College Degree	0.31 (0.01)[0.27]	
Fraction Temporary	0.05 (0.01)[0.27]	
Fraction Part Time	0.18 (0.01)[0.24]	
Fraction Covered by Collective Bargaining	0.03 (0.01)[0.15]	
Offer Health Insurance Plan	0.82 (0.02)[0.38]	
Training Department	0.19 (0.02)[0.4]	
Months to Full Productivity, New Professional	2.63 (0.15)[2.85]	
Months to Full Productivity, New Non-Professional	2.01 (0.12)[2.48]	
Months to Reach Full Group Productivity, New Professional	3.25 (0.2)[3.73]	
Months to Reach Full Group Productivity, New Non-Professional	2.05 (0.16)[3.42]	
Require a Written Test	0.3 (0.02)[0.46]	
Recruit on Internet	0.52 (0.02)[0.5]	
Post Help Wanted Signs	0.29 (0.02)[0.45]	
Importance of Seniority in Promotion	2.18 (0.05)[1.01]	
Measures of Replacement Costs Per Employee		
Replacement Costs Per Employee	439.17 (50.5)[1201.3]	4.41 (0.1)[2.2]
Replacement Costs Per Employee (1 Yr Rate)	632.22 (78.72)[2035.44]	3.85 (0.18)[2.92]
Imputed Replacement Costs Per Employee	445.24 (45.16)[1117.59]	4.91 (0.1)[1.65]

Notes: Mean values of variables and their logs (for some variables) in the table using a pooled sample of both waves of the CES. See text for details.

Table 3: Replacement Costs Per Recruit

	Mean (Std. Error)[Std. Dev.]				
	Pooled CES	Cross Section		Balanced Panel	
		2003	2008	2003	2008
All Workers	4039 (420)[9800]	4529 (544)[10782]	3177 (655)[7712]	3917 (715)[10670]	4337 (1530)[10741]
Professional and Managerial	7051 (730)[15643]	7558 (950)[16902]	5992 (1047)[12575]	6662 (1221)[16869]	7597 (1848)[15748]
Blue Collar	1928 (475)[12290]	2341 (753)[15399]	1246 (165)[2853]	1471 (230)[2692]	1481 (269)[2472]
Clerical (2003 Only)		2587 (269)[5320]		2516 (447)[5849]	
Sales (2003 Only)		4741 (698)[16377]		4976 (932)[14823]	
Clerical and Sales (2008 Only)			2406 (380)[4786]		2550 (561)[5084]

Notes: Costs are in 2003 dollars. Standard errors are in parentheses and standard deviations are in brackets. All calculations use the sampling weights for each establishment, and do not weight by the number of workers at each establishment. Note that none of the replacement costs are strictly comparable across years, as the categories about which replacement cost questions were asked changed across waves. However, in both waves, questions were asked about replacement costs for professional/managerial workers as well as blue collar/manual labor workers; on the other hand, in 2008, clerical and sales were grouped together. "Cross Section" refers to the full data sets for each of the two waves, while "Balanced Panel" refers to the subset of establishments present in both waves. Calculations for the 2003 and 2008 cross sections use establishment weights adjusted for nonresponse for the respective year, while calculations for the 2003 and 2008 balanced panel subsets use establishment weights adjusted for nonresponse for the 2003 wave and adjusted further for nonresponse for the 2008 wave. For both the 2003 and 2008 balanced panel subsets, 2003 strata are used when accounting for stratification.

Table 4: Regression of Changes on Changes, Log of Replacement Cost Per Recruit on Importance of Seniority in Promotion

	Log of Replacement Cost Per Recruit (Std. Error)[p−value]		
Variable (Differenced)	(1)	(2)	(3)
Importance of Seniority in Promotion	−1.03 (0.31)[0.001]	−0.92 (0.32)[0.004]	−0.91 (0.39)[0.02]
Log Establishment Size		0.099 (0.24)[0.67]	−0.1 (0.27)[0.7]
Average Log Wage		2.31 (1.38)[0.095]	1.81 (1.24)[0.14]
Bargaining Coverage			0.72 (3.53)[0.84]
Fraction Professional or Managerial			2.24 (2.36)[0.34]
Fraction Blue Collar			2.77 (1.36)[0.042]
Fraction College Degree			2.63 (3.42)[0.44]
Fraction Temporary			0.95 (2.48)[0.7]
Fraction Part Time			−1.65 (2.26)[0.47]
Constant	0.076 (0.54)[0.89]	−0.023 (0.51)[0.96]	−0.2 (0.42)[0.63]
N	130	124	98

Notes: Uses only the balanced panel, i.e. establishments present in both of the 2003 and 2008 waves. Standard errors are in parentheses. p-values are in brackets. Replacement costs and wages are in 2003 dollars. Technically, the response variable is the log of one plus the replacement cost per recruit. Note that the fraction in any particular occupational category is not strictly comparable across waves, as in the 2008 wave the clerical and sales categories are combined. Moreover, the fraction in each category is an imputed value, using the fractions for every category in order that they sum to 1; thus, even exactly corresponding categories are not comparable across waves. Weights used are 2003 establishment weights adjusted for nonresponse.

Table 5: Regression of Log of Replacement Cost Per Employee on Log Establishment Size and Log Wage, Pooled Sample

	Log of Replacement Cost Per Employee (Std. Error)[$p-$value]			
Variable	(1)	(2)	(3)	(4)
Log Establishment Size	0.37 (0.11)[0.001]	0.39 (0.12)[0.001]	0.36 (0.12)[0.004]	0.39 (0.12)[0.001]
Average Log Wage		1.43 (0.3)[0]	1.12 (0.36)[0.002]	1.15 (0.36)[0.001]
Bargaining Coverage			-0.3 (0.61)[0.61]	-0.085 (0.61)[0.89]
Fraction Professional or Managerial			0.071 (0.84)[0.93]	0.63 (0.8)[0.43]
Fraction Blue Collar			-0.54 (0.54)[0.32]	-0.12 (0.57)[0.84]
Fraction College Degree			-0.47 (0.8)[0.55]	-0.031 (0.77)[0.97]
Fraction Temporary			0.14 (0.21)[0.5]	0.1 (0.2)[0.62]
Fraction Part Time			-0.9 (0.62)[0.15]	-0.88 (0.57)[0.12]
Offer Health Insurance Plan			0.71 (0.34)[0.04]	0.67 (0.34)[0.05]
Constant	3.3 (0.42)[0]	-0.64 (0.86)[0.46]	0.14 (1.14)[0.9]	
N	1013	1000	879	879
One Digit Industry Fixed Effects	No	No	No	Yes

Notes: Pools all observations in 2003 and 2008 waves; establishments in both waves are treated as distinct observations. Standard errors are in parentheses, and are clustered at the level of the establishment; i.e. each cluster consists either of two observations, namely the same establishment observed in both waves, or one observation. p-values are in brackets. A $p-$value of 0 indicates that $p < 0.0005$ holds. Replacement costs per employee are estimated using the replacement cost per recruit and multiplying by the ratio of employees at the establishment less than five years to the establishment size. Replacement costs and wages are in 2003 dollars. Technically, the response variable is the log of one plus the replacement cost per employee. Note that the fraction in any particular occupational category is not strictly comparable across waves, as in the 2008 wave the clerical and sales categories are combined. Moreover, the fraction in each category is an imputed value, using the fractions for every category in order that they sum to 1; thus, even exactly corresponding categories are not comparable across waves. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.

Table 6: Regression of Changes on Changes, Log of Replacement Cost Per Employee

	Log of Replacement Cost Per Employee (Std. Error)[p−value]		
Variable (Differenced)	(1)	(2)	(3)
Log Establishment Size	−0.083 (0.21)[0.69]	−0.058 (0.19)[0.76]	−0.27 (0.2)[0.18]
Average Log Wage		1.84 (1.24)[0.14]	1.15 (1.16)[0.32]
Bargaining Coverage			−7.44 (2.85)[0.009]
Fraction Professional or Managerial			2.19 (1.92)[0.25]
Fraction Blue Collar			2.54 (1.57)[0.11]
Fraction College Degree			1.68 (2.96)[0.57]
Fraction Temporary			0.32 (1.87)[0.87]
Fraction Part Time			−3.33 (2.45)[0.17]
Constant	−0.11 (0.5)[0.83]	−0.18 (0.46)[0.7]	−0.091 (0.37)[0.81]
N	117	115	93

Notes: Uses only the balanced panel, i.e. establishments present in both of the 2003 and 2008 waves. Standard errors are in parentheses. p-values are in brackets. Replacement costs and wages are in 2003 dollars. Technically, the response variable is the log of one plus the replacement cost per employee. Note that the fraction in any particular occupational category is not strictly comparable across waves, as in the 2008 wave, the clerical and sales categories are combined. Moreover, the fraction in each category is an imputed value, using the fractions for every category in order that they sum to 1; thus, even exactly corresponding categories are not comparable across waves. Weights used are 2003 establishment weights adjusted for nonresponse.

Table 7: Regression of Changes on Changes, Log of Replacement Cost Per Employee, Checking Robustness to Outliers

	Log of Replacement Cost Per Employee (Std. Error)[$p-$value]		
Variable (Differenced)	(1)	(2)	(3)
Log Establishment Size	−0.1 (0.18)[0.59]	−0.09 (0.19)[0.63]	−0.12 (0.21)[0.57]
Average Log Wage	0.18 (1.84)[0.92]	−1.15 (1.41)[0.42]	−0.93 (1.73)[0.59]
Bargaining Coverage			−5.28 (2.92)[0.07]
Fraction Professional or Managerial			−0.73 (1.77)[0.68]
Fraction Blue Collar			−0.19 (1.19)[0.87]
Fraction College Degree			0.61 (2.55)[0.81]
Fraction Temporary			0.63 (1.25)[0.61]
Fraction Part Time			−0.79 (1.83)[0.67]
Constant	0.2 (0.41)[0.62]	0.032 (0.38)[0.93]	0.081 (0.3)[0.79]
N	110	109	87

Notes: Uses only the balanced panel, i.e. establishments present in both of the 2003 and 2008 waves. Standard errors are in parentheses. p-values are in brackets. See text for criteria for removing observations. Replacement costs and wages are in 2003 dollars. Technically, the response variable is the log of one plus the replacement cost per employee. Note that the fraction in any particular occupational category is not strictly comparable across waves, as in the 2008 wave, the clerical and sales categories are combined. Moreover, the fraction in each category is an imputed value, using the fractions for every category in order that they sum to 1; thus, even exactly corresponding categories are not comparable across waves. Weights used are 2003 establishment weights adjusted for nonresponse.

Table 8: Regression of Changes on Changes, Log of Replacement Cost Per Employee, Establishments whose Most Common Occupational Category is Blue Collar Workers

	Log of Replacement Cost Per Employee (Std. Error)[$p-$value]					
Variable (Differenced)	(1)	(2)	(3)	(4)	(5)	(6)
Log Number Blue Collar	0.18 (0.36)[0.62]	0.14 (0.34)[0.67]	0.18 (0.23)[0.42]			
Log Establishment Size				0.087 (0.31)[0.78]	0.026 (0.29)[0.93]	0.17 (0.22)[0.43]
Log Entry Pay		−0.54 (1.86)[0.77]	−0.0023 (1.57)[1]		−0.59 (1.93)[0.76]	0.011 (1.58)[0.99]
Bargaining Coverage			0.66 (3.69)[0.86]			0.73 (3.65)[0.84]
Fraction Professional or Managerial			8.78 (2)[0]			8.79 (2.01)[0]
Fraction Blue Collar			5.45 (1.49)[0]			5.71 (1.61)[0]
Fraction College Degree			−3.1 (2.65)[0.24]			−3.06 (2.64)[0.25]
Fraction Temporary			−1.92 (1.58)[0.22]			−1.9 (1.57)[0.23]
Fraction Part Time			1.06 (1.79)[0.55]			1.07 (1.8)[0.55]
Constant	−0.26 (0.25)[0.29]	−0.25 (0.23)[0.27]	−0.77 (0.24)[0.002]	−0.27 (0.24)[0.28]	−0.26 (0.23)[0.25]	−0.77 (0.24)[0.002]
N	103	100	88	103	100	88

Notes: Uses only the balanced panel, i.e. establishments present in both of the 2003 and 2008 waves. Standard errors are in parentheses. p-values are in brackets. A $p-$value of 0 indicates that $p < 0.0005$ holds. Replacement costs and wages are in 2003 dollars. Technically, the response variable is the log of one plus the replacement cost per employee for blue collar worker and manual laborers. Restricted to establishments reporting that blue collar employees are their most common occupational category in both waves. Note that the fraction in any particular occupational category is not strictly comparable across waves, as in the 2008 wave, the clerical and sales categories are combined. Moreover, the fraction in each category is an imputed value, using the fractions for every category in order that they sum to 1; thus, even exactly corresponding categories are not comparable across waves. Weights used are 2003 establishment weights adjusted for nonresponse.

Table 9: Regression of Changes on Changes, Log of Replacement Cost Net of Training Costs Per Employee

	Log of Rep. Costs Net Train. Costs Per Emp. (Std. Error)[p−value]		
Variable (Differenced)	(1)	(2)	(3)
Log Establishment Size	−0.14 (0.14)[0.31]	−0.16 (0.14)[0.26]	−0.33 (0.24)[0.17]
Average Log Wage		−0.42 (0.96)[0.66]	−1.26 (0.99)[0.2]
Bargaining Coverage			−11 (3.84)[0.004]
Fraction Professional or Managerial			0.07 (2.11)[0.97]
Fraction Blue Collar			0.8 (1.38)[0.56]
Fraction College Degree			1.93 (2.7)[0.48]
Fraction Temporary			−2.35 (2.22)[0.29]
Fraction Part Time			−4.59 (2.76)[0.097]
Constant	0.064 (0.4)[0.87]	0.088 (0.38)[0.82]	0.18 (0.4)[0.65]
N	103	101	81

Notes: Uses only the balanced panel, i.e. establishments present in both of the 2003 and 2008 waves. Standard errors are in parentheses. p-values are in brackets. Replacement costs and wages are in 2003 dollars. Technically, the response variable is the log of one plus the replacement cost per employee. Note that the fraction in any particular occupational category is not strictly comparable across waves, as in the 2008 wave, the clerical and sales categories are combined. Moreover, the fraction in each category is an imputed value, using the fractions for every category in order that they sum to 1; thus, even exactly corresponding categories are not comparable across waves. Weights used are 2003 establishment weights adjusted for nonresponse.

Table 10: Regression of Log of Replacement Cost per Recruit on Log Establishment Size and Log Wage

	Log of Replacement Cost Per Recruit (Std. Error)[$p-$value]			
Variable	(1)	(2)	(3)	(4)
Log Establishment Size	0.58 (0.14)[0]	0.61 (0.15)[0]	0.51 (0.16)[0.001]	0.53 (0.15)[0]
Average Log Wage		1.79 (0.4)[0]	1.52 (0.49)[0.002]	1.44 (0.46)[0.002]
Bargaining Coverage			-0.11 (0.77)[0.88]	0.23 (0.77)[0.76]
Fraction Professional or Managerial			0.32 (1.1)[0.77]	1.03 (1.02)[0.31]
Fraction Blue Collar			-0.077 (0.72)[0.91]	0.26 (0.75)[0.73]
Fraction College Degree			-1.08 (1.08)[0.32]	-0.46 (1.04)[0.66]
Fraction Temporary			0.23 (0.24)[0.34]	0.21 (0.23)[0.36]
Fraction Part Time			-1.42 (0.74)[0.055]	-1.27 (0.67)[0.059]
Offer Health Insurance Plan			0.99 (0.47)[0.034]	0.88 (0.45)[0.047]
Constant	4.68 (0.54)[0]	-0.26 (1.17)[0.82]	0.42 (1.52)[0.78]	
N	1013	1000	879	879
One Digit Industry Fixed Effects	No	No	No	Yes

Notes: Pools all observations in either 2003 or 2008 wave; establishments in both waves are treated as distinct observations. Standard errors are in parentheses, and are clustered at the level of the establishment; i.e. each cluster consists either of two observations, namely the same establishment observed in both waves, or one observation. p-values are in brackets. A $p-$value of 0 indicates that $p < 0.0005$ holds. Replacement costs and wages are in 2003 dollars. Note that the fraction in any particular occupational category is not strictly comparable across waves, as in the 2008 wave the clerical and sales categories are combined. Moreover, the fraction in each category is an imputed value, using the fractions for every category in order that they sum to 1; thus, even exactly corresponding categories are not comparable across waves. One observation, which seems to be erroneously identified as being in the government one-digit sector, is dropped. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.

Table 11: Regression of Log of Recruitment Rate on Log Establishment Size and Log Wage

	Log of Recruits Per Employee (Std. Error)[$p-$value]			
Variable	(1)	(2)	(3)	(4)
Log Establishment Size	-0.0088 (0.0029)[0.002]	-0.0092 (0.0029)[0.001]	-0.0073 (0.0031)[0.02]	-0.0063 (0.0032)[0.048]
Average Log Wage		-0.021 (0.0075)[0.006]	-0.014 (0.0099)[0.15]	-0.013 (0.011)[0.24]
Bargaining Coverage			-0.034 (0.014)[0.014]	-0.037 (0.014)[0.01]
Fraction Professional or Managerial			-0.037 (0.017)[0.035]	-0.027 (0.019)[0.15]
Fraction Blue Collar			-0.024 (0.013)[0.068]	-0.018 (0.014)[0.21]
Fraction College Degree			0.027 (0.018)[0.14]	0.029 (0.019)[0.12]
Fraction Temporary			0.0003 (0.007)[0.96]	-0.0013 (0.0068)[0.85]
Fraction Part Time			0.037 (0.019)[0.046]	0.034 (0.018)[0.062]
Offer Health Insurance Plan			-0.013 (0.0095)[0.17]	-0.0093 (0.0096)[0.33]
Constant	0.14 (0.011)[0]	0.2 (0.024)[0]	0.19 (0.032)[0]	
N	1013	1000	879	879
One Digit Industry Fixed Effects	No	No	No	Yes

Notes: Pools all observations in either 2003 or 2008 wave; establishments in both waves are treated as distinct observations. Standard errors are in parentheses, and are clustered at the level of the establishment; i.e. each cluster consists either of two observations, namely the same establishment observed in both waves, or one observation. p-values are in brackets. A $p-$value of 0 indicates that $p < 0.0005$ holds. Wages are in 2003 dollars. Note that the fraction in any particular occupational category is not strictly comparable across waves, as in the 2008 wave the clerical and sales categories are combined. Moreover, the fraction in each category is an imputed value, using the fractions for every category in order that they sum to 1; thus, even exactly corresponding categories are not comparable across waves. One observation, which seems to be erroneously identified as being in the government one-digit sector, is dropped. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.