Essays in Labor Economics

https://escholarship.org/uc/item/0210k3c9

Freeman, Donald Eric

2010

Peer reviewed|Thesis/dissertation
Essays in Labor Economics

by

Donald Eric Freeman

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Economics in the Graduate Division of the University of California, Berkeley

Committee in charge:
Professor Michael Reich, Chair
Professor Clair Brown
Professor Steven Raphael

Spring 2010
Abstract

Essays in Labor Economics

by

Donald Eric Freeman
Doctor of Philosophy in Economics
University of California, Berkeley
Professor Michael Reich, Chair

We first 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 (although not robust) relationship between replacement costs and the wage. This result constitutes an anomaly for hiring and separation models, such as Manning (2003), in which the negative wage elasticity of replacement costs is a key assumption. We also examine particular occupational groups, and find that although there is not necessarily a clear incongruity for blue collar workers, there is one for professional and managerial workers. To resolve the anomaly, we propose several possible solutions including for one the role of heterogeneity, and on the other hand expanding Manning’s model to incorporate the effects of wages on productivity.

We also study the relationship of wages with High Performance Workplace Organization (HPWO) practices, a broadly-defined set of changes in establishment policies, often involving flexible policies, that have been adopted by many firms in the U.S. in the past few decades. HPWO practices include, for example, employee meetings to discuss workplace problems, self-managed teams, and job rotation. There is no empirical consensus in the literature on the relationship between wages and these practices, and the data required for its study is somewhat uncommon. We use the same panel survey of California businesses to investigate the issue. We find some evidence that adoption of these practices is correlated with increased entry wages for blue collar workers, and correlated with increased entry and highest tier wages for professionals and managers. There is also evidence that HPWO incidence is associated with increased wages for workers at manufacturing establishments, establishments with few professional or managerial workers, and workplaces with a union presence. Regarding individual practices, we find support for a positive correlation between
higher wages and increased worker participation in self-managed teams, especially among blue collar workers and at establishments with few professional or managerial workers. But there is a negative relationship between team adoption and wages for professionals and managers. Finally, the adoption of formal quality management programs is associated with higher wages for professionals and managers. We conclude that the relationship between HPWO intensity and wages is complex and heterogeneous, varying for different groups of workers and different practices, and in fact some of these correlations may lead to cancelling effects when one looks only at aggregate relationships. Thus any analysis of high performance practices should be examined with a focus on fine levels of interaction.

We also study the interplay among tenure, firm size and firm age. The positive association between firm size and the average tenure of employees at a firm is well-known. We present evidence, at the establishment level, that a large portion, though likely not all, of this association can be explained by the following observation: there is a positive association between an establishment’s size and its age so that when combined with the mechanical relationship between establishment age and average employee tenure, this fact leads to a size-tenure correlation. We investigate this issue by making use of the panel survey of California establishments; it is very useful for studying the topic because it is unusual in having data on all three of these quantities.

We also study the relationship of certain employer policies with voluntary quit rates, using the panel survey of California businesses. We particularly focus on two policies, namely (i) how much employers give preference to current employees when hiring for a position above entry level, and (ii) how much employers take into account seniority when promoting to a position they have already decided to fill internally. This question has been studied before by Fairris (2004), but the data required is scarce; moreover, with this survey of California establishments, we can answer new related questions. We find some evidence from pooled regressions that (i) is associated positively with quit rates. Balanced panel regressions do not confirm these results; however, the standard errors are quite large due to small sample sizes. We find no clear evidence of associations between (ii) and the overall voluntary quit rate. Moreover, the data set allows us to investigate questions that have not yet been studied, by breaking down quits into finer categories. We find, albeit with non-robust results, that (i) is positively associated in particular with quits to accept another job. We find as well that (i) and (ii) are both negatively associated with employee turnover rates due to unsatisfactory performance.

We also examine a controlled randomized health care experiment known as the RAND Health Insurance Experiment, which tested the effect of free insurance on use of medical services and on health outcomes, relative to control groups which paid a larger share of health costs out-of-pocket. We focus on using the experiment to test econometric techniques, by comparing experimental estimates of the benefit of free insurance on certain health outcomes to estimates one might obtain using nonexperimental econometric techniques if an analogous “natural experiment” were
all that was available, following Lalonde (1986). Like Lalonde, our results cast doubt on the use of these nonexperimental methods. We then present a closer inspection of both the setup of the experiment and some findings of previous researchers. This work yields a slight caveat in the form of one piece of evidence that the result we have used in the test may not be robust. We also consider other health outcomes, and identify one beneficial impact of free health insurance that does not appear to have been noticed previously by RAND researchers, as well as finding several apparently unnoticed beneficial effects that vary by income category. Like the RAND researchers, we do not find very many beneficial impacts of the free plans, but when interpreting the experimental results, one should bear in mind the low maximums on total out-of-pocket expenditures under the control group insurance plans.
Contents

List of Figures iv
List of Tables v

1 Employee Replacement Costs 1
  1.1 Introduction ................................................. 1
  1.2 Related Literature ........................................ 4
    1.2.1 Theory .............................................. 4
    1.2.2 Empirical Work ...................................... 5
  1.3 Data ......................................................... 6
  1.4 Replacement Costs per Recruit ........................... 8
    1.4.1 Descriptive Summary Statistics ....................... 9
    1.4.2 Detailed Descriptive Statistics ................... 11
    1.4.3 Replacement Costs and Seniority in Promotion Decisions . 14
  1.5 Replacement Costs per Employee ......................... 16
    1.5.1 Variable Definitions ................................ 16
    1.5.2 Graphical Analysis .................................. 17
    1.5.3 Regression Analysis .................................. 17
  1.6 Interpretation ............................................. 22
    1.6.1 Labor Cost Model .................................... 22
    1.6.2 Interpretation of Results in Context of the Manning Labor Cost Model ......................... 25
    1.6.3 Discussion of Results on Wage Elasticity .......... 26
  1.7 Conclusion .................................................. 32
  1.8 Data Appendix ................................................ 34
    1.8.1 Wage Measures ....................................... 34
    1.8.2 Imputed Replacement Costs ......................... 34

2 Wage Effects of High Performance Workplace Practices 60
  2.1 Introduction ................................................ 60
  2.2 High Performance Workplace Practices: Theory and Related Literature 61
  2.3 Data .......................................................... 64
5.5.1 Comparing Non-experimental Evaluation with Experimental Evaluation ........................................ 149
5.5.2 Validity of the RAND Experiment and Findings ................................................................. 153
5.5.3 Other Outcomes of the RAND Experiment ..................................................................... 156
5.6 Conclusion ......................................................................................................................... 159

Bibliography .......................................................................................................................... 174
List of Figures

1.1 Mean Replacement Costs (RCR) ........................................ 36
1.2 A Simple Labor Market Model with a Fixed Wage and Increasing Replacement Costs .................................................. 37
1.3 The Classical Static Monopsony Model ............................. 38
1.4 Mean RCR, by Survey Wave ............................................. 39
1.5 Histograms of RCR (See text for details.) ......................... 40
1.6 Mean and Median of RCR to Annual Wage ......................... 41
1.7 Histogram of Training Hours for New Employees, Pooled Sample ................................. 42
1.8 Histograms of Time to Full Productivity for New Employees ................................. 43
1.9 Scatter Plot of Log of RCE by Log of Establishment Size .......... 44
1.10 Scatter Plot of Log of RCE by Log Wage ............................ 45
1.11 The Labor Cost Function for the Basic Competitive Model .... 46
1.12 The Labor Cost Function for the Basic Monopsony Model .... 46
2.1 Frequencies of Responses for Individual HPWO Practices .......... 82
3.1 Average Tenure by Establishment Size Category .................. 108
List of Tables

1.1 Estimates of Replacement Costs (RCR, Dollar Value or Percentage of Total Labor Costs) ................................................. 47
1.2 Sampling Results and Response Rates in the CBS .................. 48
1.3 Descriptive Statistics for Key Variables, Pooled Sample .......... 49
1.4 Replacement Costs Per Recruit (RCR) ............................... 50
1.5 Regression of Changes on Changes, Log of RCR on Importance of Seniority in Promotion .............................................. 51
1.6 Regression of Log of RCE on Log Establishment Size and Log Wage, Pooled Sample .................................................. 52
1.7 Regression of Log of RCE for Professionals and Managers on Log Establishment Size and Log Wage of Most Recent Professional or Managerial Hire, Pooled Sample .............................................. 53
1.8 Regression of Log of RCE for Blue Collar Workers on Log Establishment Size and Log Wage of Most Recent Nonexempt Hire, Pooled Sample .................................................. 54
1.9 Regression of Changes on Changes, Log of RCE on Log Establishment Size and Log Wage .............................................. 55
1.10 Regression of Log of RCE on Log Establishment Size and Log Wage, Pooled Sample, Effects of Time to Full Productivity and New Training Hours .................................................. 56
1.11 Regression of Changes on Changes, Log of Replacement Costs Net of Production Value Per Employee .......................... 57
1.12 Regression of Log of RCR on Log Establishment Size and Log Wage .................................................. 58
1.13 Regression of Log of Recruitment Rate on Log Establishment Size and Log Wage .................................................. 59
2.1 Sampling Results and Response Rates in the CBS .................. 83
2.2 Definitions of Key Variables ............................................ 84
2.3 Descriptive Statistics for Key Variables (Pooled Sample) .......... 85
2.4 Regression of Log Hourly Entry Wage for Blue Collar Employees on Summary HPWO Measure, Pooled Sample, Only Establishments Whose Most Common Occupation is Blue Collar .................. 86
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3 Regression of Changes in Average Tenure on Changes in Establishment Size and Age and Other Variables, Balanced Panel</td>
<td>111</td>
</tr>
<tr>
<td>4.1 Sampling Results and Response Rates in the CBS</td>
<td>128</td>
</tr>
<tr>
<td>4.2 Definitions of Key Variables</td>
<td>129</td>
</tr>
<tr>
<td>4.3 Descriptive Statistics for Key Variables (Pooled Sample)</td>
<td>130</td>
</tr>
<tr>
<td>4.4 Regression of Quit Rate on Health Insurance Eligibility for Full Time Employees, Pooled Sample</td>
<td>131</td>
</tr>
<tr>
<td>4.5 Regression of Quit Rate on Various Internal Labor Market Policies, Pooled Sample</td>
<td>132</td>
</tr>
<tr>
<td>4.6 Regression of Quit Rate due to Accepting Another Job on Various Internal Labor Market Policies, 2003 Wave Only</td>
<td>133</td>
</tr>
<tr>
<td>4.7 Regression of Turnover Rate due to Unsatisfactory Performance on Various Internal Labor Market Policies, 2003 Wave Only</td>
<td>134</td>
</tr>
<tr>
<td>4.8 Regression of Quit Rate due to Accepting Another Job on Monthly Employee Copayment, 2003 Wave Only</td>
<td>135</td>
</tr>
<tr>
<td>5.1 Descriptions of Selected RAND HIE Variables</td>
<td>162</td>
</tr>
<tr>
<td>5.2 Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Diastolic Blood Pressure on Treatment Indicator and Various Controls</td>
<td>163</td>
</tr>
<tr>
<td>5.3 Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Systolic Blood Pressure on Treatment Indicator and Various Controls</td>
<td>164</td>
</tr>
<tr>
<td>5.4 Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Diastolic Blood Pressure on Treatment Indicator and Different Sets of Controls, using Heckman Selection Correction</td>
<td>165</td>
</tr>
<tr>
<td>5.5 Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Systolic Blood Pressure on Treatment Indicator and Different Sets of Controls, using Heckman Selection Correction</td>
<td>166</td>
</tr>
<tr>
<td>5.6 Comparison of Means of Various Characteristics and Health Measures at Entry, and Parametric Two-Sided t—tests of Equality of Means</td>
<td>167</td>
</tr>
<tr>
<td>5.7 Comparison of Means of Additional Health Measures at Entry, and Parametric Two-Sided t—tests of Equality of Means</td>
<td>168</td>
</tr>
<tr>
<td>5.8 Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Functional Far Vision Measure at Exit on Treatment Indicator and Various Sets of Regressors, using Two Control Groups</td>
<td>169</td>
</tr>
<tr>
<td>5.9 Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Differences in Outcome Measures between Entry and Exit, on Treatment Indicator and Various Sets of Regressors, using two Control Groups</td>
<td>170</td>
</tr>
</tbody>
</table>
5.10 Comparison of Mean Outcomes at Exit, and Parametric One-Sided $t$–tests of Lack of Change ........................................... 171
5.11 Comparison of Mean Outcomes at Exit, and Parametric One-Sided $t$–tests of Lack of Change, Within Subpopulations Categorized by Family Size .................................................. 172
5.12 Comparison of Mean Outcomes at Exit, and Parametric One-Sided $t$–tests of Lack of Change, Within Subpopulations Categorized by Income .................................................. 173
Acknowledgments

There are numerous people to whom I owe gratitude— I unfortunately cannot mention them all. First, I would like to thank David Fairris for helpful comments and Paul Osterman for providing access to the instrument for his survey. I would also like to thank the superb and exceptionally kind staff at IRLE and the Economics Department, especially Patrick Allen. I thank Suresh Naidu, Arin Dube, Ethan Kaplan, and Sylvia Allegretto for their excellent suggestions, extremely helpful advice and encouragement. I thank Steven Raphael for helpful comments and generosity with his time. I am especially grateful to Clair Brown for her numerous careful and insightful suggestions and incredible effort to help me. I owe a phenomenal debt to my advisor Michael Reich for encouragement and unyielding support throughout my graduate career. I also thank Arin Dube and Michael Reich, in their capacity as coauthors, for allowing me to use our joint work in the dissertation.

On the nonacademic side, I would like to thank, among many friends, Brendan Weickert, Teddy Keizer, Scott Robinson, Kirsten Green and Dave Graham-Squire, for encouragement throughout the process. I also would like to thank my brothers Tim and Greg, and my parents for all their support, and especially my mother for all of her encouragement and advice. And finally, for everything, I simply cannot say enough to thank Julia.
Chapter 1

Employee Replacement Costs

1.1 Introduction

Although wages constitute a core topic in empirical and theoretical labor economics research, the nonwage labor costs to firms of replacing a worker have been much less investigated. In part, this lacuna results from the scarcity of representative microdata that contain both information on the cost of replacing an employee and on characteristics of the workplace. As Manning (2006, p. 90) notes, “since Oi’s pioneering 1962 study which uses data from 1951 there have been hardly any papers written on the size of turnover costs.” Yet such costs may be substantial enough to affect decisions on hiring and employee retention policies.

The purpose of this paper is to investigate the economics of employee replacement costs, using a new panel survey of California workplaces, the California Business Survey (CBS). We define the cost of replacing a worker broadly, to include recruitment, selection and screening costs, the costs of learning on the job and becoming fully productive, and separation costs. The CBS is distinctive in combining a comprehensive set of measures of these replacement costs with detailed information on workplace characteristics and practices.

One focus of this paper is to report the levels and distributions of these replacement costs and their relationships with other variables. Figure 1 reports the mean, across establishments, of the average replacement costs for all workers 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.

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. 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.

These costs constitute a significant fraction of annual wage costs and presumably affect their personnel decisions. For example, the expenditures could incentivize firms to try to lower quits by increasing pay.

The components of replacement costs are also large. For example, the mean establishment trains new workers for 65.6 hours. If some of this training is general rather than firm-specific, then an interesting consequence, as observed by Acemoglu and Pischke (1999), is that firms must actually have monopsony power; after all, if employers could not obtain rents from workers by paying a wage below worker productivity, they could not recover their investment in training.\footnote{Even if the training were all firm-specific, this could be suggestive of a monopsony model if the firm does not \textit{fully} capture all of the returns from training. To see this, observe that in the standard discussion of firm-specific capital, Becker (1964) also gives an intuitive argument why a firm might share rents in order to deter quits. The standard formalization of this argument, given by Hashimoto (1981), essentially assumes that the labor supply curve to the firm is not infinitely elastic, which is the key feature of a monopsonistic labor market. This observation is due to Manning (2003); see pp. 129-131 for details.}

As mentioned earlier, we focus not only on estimates of replacement costs, but also on their relationship with multiple quantities. For example, there is reason to believe hiring and promotion policies have an important effect on replacement costs. Consider one such policy, the use of seniority in promotion.\footnote{This variable measures the priority a firm places on seniority in a promotion decision, given that the firm has already decided to fill a particular vacancy from within.} By choosing to hire from within in the first place, a firm incurs costs from both the promotion and from an external hire to fill the vacated lower level position. If the position vacated depends on how the firm makes promotion decisions, then the seniority measure could affect the costs of the external hire. But this would not be the case if promotions are made from a pool of candidates all at a given job level, which may be a common situation.

In this environment, the use of seniority could be negatively correlated with replacement costs per recruit for two reasons. For one, prioritizing higher seniority could reduce training costs if senior workers are more likely to have establishment-specific skills. Alternatively, the replacement process could be much less expensive if promotion decisions are based on the simple criterion of seniority, rather than being reached through a lengthier process involving, perhaps, multiple interviews and managerial decision-making meetings.

A key theme of the paper is the relationship between establishment size and replacement costs per employee, which is the total annual replacement costs at an establishment divided by the total number of employees. A positive correlation could be a result of a monopsonistic labor market. Consider the classical setting of monopsony (or oligopsony) in which one firm hires a large share of workers in a given labor
market. If a firm grows sufficiently large relative to the pool of labor, then rather than raising the wage, it might increase recruiting effort in an attempt to hire more workers. As this increased effort is presumably more expensive, replacement costs per worker would grow larger with establishment size. This demonstrates why, if a labor market exhibits monopsony, one might expect to see a positive relationship between establishment size and replacement costs.

Suppose now that in a particular labor market, there is evidence of a positive relationship between size and replacement costs per worker. Is there reason to believe that, reciprocally, the labor market exhibits monopsony? It is worth taking a moment first to clarify that by monopsony we mean, as usual, the situation in which the labor supply to an individual firm is not infinitely elastic, rather than solely the more narrow situation in which there is only one firm purchasing labor.

Returning to the question at hand, suppose that larger firms do in fact spend more money per extra worker. For simplicity, assume that firms are price-takers in the labor market, i.e. that there is a fixed wage \( w_0 \). For now, to simplify our informal exposition, we implicitly assume away any changes in the rate of worker quits; the later discussion of the formal model shows that this is not necessary.

Consider the resulting marginal cost of labor (MCL) curve. If there were no replacement costs, MCL would be exactly \( w_0 \), just as in the standard competitive model. But because there is a nonzero cost of recruitment per worker, the MCL is strictly larger than the wage \( w_0 \). More importantly, because the recruitment cost per worker increases in the number of workers, the MCL is actually an increasing function of the number of workers. Figure 1.2 depicts this scenario, with a flat wage schedule and an increasing MCL curve. Note the similarities and contrasts with Figure 1.3, which portrays the classical static monopsony model.

To summarize, the conclusion is that the MCL curve is upward-sloping, just as in a monopsony model. So the labor market exhibits monopsony if there is a positive correlation between establishment size and replacement costs.

This situation is formalized in a model due to Manning (2003), which we discuss later. In that model, as above, the key issue is the relationship between establishment size and replacement costs per employee, and there is evidence for monopsony if there is a positive partial correlation.

In our empirical work, we find that replacement costs per employee grow with the size of the establishment, which indicates a monopsonistic labor market. We also find evidence for a positive partial correlation between replacement costs per employee and the wage, which is at odds with the Manning model.

The plan of the paper is as follows. We first review the literature on replacement costs. We describe the California Business Survey and some of its key variables in section 1.3. Section 1.4 reports generally on replacement costs per recruit. Here we also investigate the association between replacement costs and the use of seniority in establishment size and replacement costs.

\( ^6 \)If in fact firms must raise the wage to attract more workers, then the situation of course already demonstrates monopsony.

\( ^7 \)For now, to simplify our informal exposition, we implicitly assume away any changes in the rate of worker quits; the later discussion of the formal model shows that this is not necessary.
promotion. We turn to replacement costs per employee in section 1.5, and empirically investigate their relationship both with establishment size and the wage. In section 1.6, we interpret these relationships in light of Manning’s model and we suggest an expansion of the model that is consistent with the estimated positive relationship between replacement costs per employee and the wage rate. The final section concludes.

1.2 Related Literature

1.2.1 Theory

The model of Manning (2003, 2006) discussed above plays a central role in our paper; we mention some other key theoretical contributions here. Oi (1962) gave an early model that incorporated training costs; standard short-run competitive models had assumed that all labor costs depend only on the wage. In his well-known work, Becker (1964, 1993) discussed how general and specific training could affect equilibria in a competitive labor market. Many other papers discuss training; as only one example, Acemoglu and Pischke (1998) show how an employer’s asymmetric information about its workers’ abilities can give the firm monopsony power, which can lead to more training.

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.) Acemoglu and Pischke (1999) survey treatments of training costs in particular in non-competitive frameworks. Burdett and Vishwanath (1988) propose a distinction between the concepts of random matching, in which the number of recruits applying to a firm is the same for all firms, and the notion of balanced matching, in which the number of recruits is proportional to the size of the firm. They criticize the former, noting that search models often assume it, and present a model of the latter case. We note only that if random matching is in fact important, then a larger establishment may have to expend more resources on recruiting per employee than a smaller one, to make up for this underlying disadvantage. Mortensen (2000) presents a monopsony model incorporating training costs that combines two key classes of search models. Additionally, 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.

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.

1.2.2 Empirical Work

Data on recruitment and separation costs are hard to come by, as Manning notes in the above quote, but there is some work on this issue, although much of it is dated and, further, estimates from representative surveys are especially rare. 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.” This estimate, and many other estimates in this section, are summarized in Table 1. 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.8

Hamermesh (1993) reports some estimates.9 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.10

Another more recent example (Abowd and Kramarz, 2003) examines the French labor market. Naturally, results from the French labor market should be expected to vary significantly from those for the U.S. considering the different regulatory environment, including, for example, mandatory severance payments for most separations.11 Kramarz and Michaud (2010) expand on that work, reporting numerous findings, such as larger hiring costs for indefinite contracts than for 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).12 Manning (2003) reports data from 1997 through 1999 from the United Kingdom’s Institute for Personnel Development on training and recruitment costs,

---

8The survey does not appear to be a representative sample of the U.S., but it contains data on more than 5,000 employers in 11 states. See Campbell (1993) and Barron, Bishop and Dunkelberg (1985) for more detail.
9See Hamermesh (1993, p.208). In this section and throughout the paper we use the CPI-U to adjust for inflation.
from a survey of personnel officials at firms with more than 100 employees. The public sector is over-represented, as personnel officials in that sector are more likely to be members of the Institute. He finds an average turnover cost of £4,823 for managerial and £937 for unskilled workers, or roughly $8,968 and $1,742, respectively, in 2003 dollars.\footnote{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 .}

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.\footnote{We use an exchange rate of 1.6583 Swiss francs per dollar on January 1, 2002, accessed from http://www.oanda.com/currency/historical-rates .} They also estimate the cost of hiring as a function of 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.

Finally, we note that Manning (2003, 2006) considers the question of whether there is a partial correlation between replacement costs per employee and firm size, using the United Kingdom Institute for Personnel Development data on turnover costs mentioned earlier. He finds evidence for a positive partial correlation, which is indicative of monopsony. We discuss this result in more detail below.

Summarizing the descriptive results, there are not that many sources for estimates of replacement costs, and there is particularly little recent information for the U.S. market. Moreover, there is even a wide range among the estimates of replacement costs reported, perhaps because many of these estimates do not derive from a representative survey.

### 1.3 Data

A recent data set is in fact based on a representative survey, within the state of California. This data set, called the California Business Survey or CBS (Survey Research Center, 2003 and 2008), consists of a panel survey of California establishments. The data are drawn from 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 to the two waves as the 2003 and 2008 waves. The 2008 survey data contains 358 establishments from the first wave, i.e., the CBS 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 some differences. The surveys were conducted by the Survey Research Center at the University of California at Berkeley.\textsuperscript{15}

The CBS includes data on numerous characteristics of each sampled establishment, including number of employees, pay scale and benefits, occupational distribution, employee tenure, hiring and training practices, general workplace practices and detailed information on the most recently hired non-exempt and professional employees.\textsuperscript{16}

The sampling universe is the set of all nonprofit and for profit establishments in 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.

\textit{Sampling procedure}

The CBS 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 (2,806 in 2003) was checked to ensure that the establishments met the eligibility criteria, namely having at least five employees, and excluding government, public schools and universities, and agriculture, fishing and forestry; this check of eligibility was because for some cases the D&B database was out of date or inaccurate. A total of 2,200 establishments met the eligibility criteria; of these, 1,080 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.\textsuperscript{17}

In addition to attempting repeatedly to contact all of the 2003 survey participants,\textsuperscript{15}\textsuperscript{16}\textsuperscript{17}

\textsuperscript{15}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.

\textsuperscript{16}As pointed out by Osterman (1994), when surveying establishments, rather than firms, a respondent may be more knowledgeable of workplace-level policies and practices, and so the reported data are more likely to be accurate. In the theoretical discussion throughout the paper, we may refer to either firms or establishments or workplaces. In the empirical work, we focus exclusively on establishments (workplaces), the level of observation in our data.

\textsuperscript{17}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.
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 1.2 gives a summary of response rates for the sample. Establishments were split into seven 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+. 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 1.2 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.

All weights that we use account for the rate of oversampling by stratum and as well account for nonresponse. When individual workers are the focus of our inquiry, we adjust as well for the number of workers in an establishment. 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.

To check the validity of the sample, we compared industry and wage distributions in the CBS to the respective distributions in the Current Population Surveys for 2003 and 2008. The distributions for each wave were very close.

### 1.4 Replacement Costs per Recruit

In section 1.4.1, we describe the data on replacement costs available from the CBS, and the summary measure we use, before reporting some descriptive statistics and basic graphs in section 1.4.2. We focus here on replacement costs per recruit, i.e. the costs of replacing one worker during one recruitment event. In section 1.4.3,
we investigate some relationships between replacement costs per recruit and other quantities.

1.4.1 Descriptive Summary Statistics

A question in the CBS 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.

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 cost(s) per recruit, which we denote by RCR (RCRs).

Table 1.3 provides descriptive statistics for both waves of the CBS, 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. The calculations account for the stratification, and also cluster at the establishment level, to account for establishments present in both waves.

In Table 1.3, the first column gives the mean of the variable, while the second column gives the mean of the log of the variable. The average RCR is $4,039. The standard deviation is quite large, at $9,800. The average of the log of the replacement cost per recruit is 6.41. The mean establishment employs 44 workers.

The CBS does not directly ask about the number of recruits or hires in the last year. The first recruits variables in Table 1.3, computed using what we call the 1 year rate, estimates this number by multiplying the fraction of full-time workers who have

23Standard errors are in parentheses while standard deviations are in brackets.
24Note 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.)
25There is one exception, which we note shortly.
26For 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.
been at the establishment for less than one year by the number of employees at the workplace.\textsuperscript{27}

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 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.\textsuperscript{28}

The entry wage is the hourly 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 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.\textsuperscript{29} (Recall that the CBS does not include government establishments.)

\textsuperscript{27}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.

\textsuperscript{28}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.

\textsuperscript{29}Our worker-weighted estimates are 12.3 percent in 2003 and 8.8 percent in 2008.
The CBS 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.\textsuperscript{30} 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 productivity after a new hire, again for each of the occupational categories mentioned above.\textsuperscript{31} Although not reported in the table, 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.\textsuperscript{32}

1.4.2 Detailed Descriptive Statistics

In this section, we look at replacement costs (per recruit) in more detail. Figure 1.1 presented their means, pooled across the 2003 and 2008 surveys. Here we consider these costs separately for each wave. Figure 1.4 gives a summary measure for all workers at each establishment, by occupational group and by survey wave.\textsuperscript{33} Table 1.4 reports the average RCR for each occupational category for the cross-sections in each year (denoted as 2003 or 2008 Cross Section), and also the RCR 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

\textsuperscript{30}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 manageral 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).

\textsuperscript{31}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.

\textsuperscript{32}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.

\textsuperscript{33}These results are means across establishments. The line segments at the top of each bar represents a 95 percent confidence interval for the mean.
in cost across years using the variation within establishments.\textsuperscript{34} The results, with standard errors and standard deviations, are reported in Table 1.4.

The average RCR 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 RCRs 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 1.4, we can compare costs for all workers, as well as for the three occupational categories of professional and managerial workers, blue collar and manual labor employees, and the combined category of clerical and sales workers.\textsuperscript{35} We see an increase in (real) RCRs for all workers, and for professional and managerial workers, but a decrease for clerical and sales workers, and a very small increase for blue collar employees.\textsuperscript{36}

These results are somewhat at odds with the overall decrease in RCR 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 RCRs, as we will see later.\textsuperscript{37} 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 1.5 presents histograms of RCRs, 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.\textsuperscript{38} Estimates of the densities use the sampling weights. Although there are outlying observations as large as $300,000, we only depict the subset of costs less than $25,000.

The top graph in Figure 1.5 shows that most establishments have a RCR (averaged over all employees) lower than $10,000. There is a fairly long right tail; 2.5 percent of establishments have a RCR 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 RCRs for professionals and managers of more than $25,000. Finally, for blue collar and manual labor employees, the distribution is

\textsuperscript{34}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.

\textsuperscript{35}Note that the last category is an imputed estimate for 2003, using relative fractions of workers in each of the sales and clerical categories.

\textsuperscript{36}The sample sizes are, respectively, 130, 161, and 163 for all workers, professional/managerial employees and blue collar employees.

\textsuperscript{37}The nonresponse could be either at the level of the observation (i.e., unit nonresponse) or the question (i.e., item nonresponse).

\textsuperscript{38}As for the other questions, clerical and sales workers were treated as distinct in 2003, but combined into one group in 2008.
much more bunched near 0, with a large spike in the $0 to $500 bin, while only 0.5 percent of establishments have average RCR greater than $25,000 for this category of employees. The histograms demonstrate the large amount of variation in RCR in general and particularly the long right tail of the distribution; moreover, RCRs for blue collar workers are unsurprisingly much more concentrated near $0 than costs for professional workers.

We examine next the ratio of RCR to the average annual wage for workers.\textsuperscript{39} Figure 1.6 shows the mean and median of this establishment-level variable, by survey wave; the mean was reported earlier for the pooled sample.

In Figure 1.6, 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). 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.\textsuperscript{40} Figure 1.7 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.\textsuperscript{41} In the top two graphs in Figure 1.8, we give histograms for the pooled sample, for each category.\textsuperscript{42} The $y$–axis here measures the fraction of observations, as opposed to the density. To improve readability, we do not show observations

\textsuperscript{39}To be clear, we now return to discussion of all workers in the establishment, not simply particular occupations.

\textsuperscript{40}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.

\textsuperscript{41}Recall that we noted above that neither is strictly consistent across waves, but particularly the latter.

\textsuperscript{42}Respondents report ranges of values; we represent these by the midpoints of the ranges.
with responses greater than 12 months. 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, which we have represented as .5, was offered as a response in the 2008 wave but not in the 2003 wave; this 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, but otherwise, they are hard to compare due to the different ranges used across waves in the questions.

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.

### 1.4.3 Replacement Costs and Seniority in Promotion Decisions

As discussed above, there are reasons to believe the use of seniority in promotion is negatively correlated with replacement costs per recruit. The CBS 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. 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 1, is 2.18.

To test the seniority hypothesis, we regress the 2003 to 2008 difference in the log of the average RCR 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 1 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 –

---

43 There are 3 such responses for professionals, and 2 for non-professionals.
44 See notes for Figure 1.8 for details of the ranges.
45 Fairris (2004) found a negative relationship between the importance of seniority in promotion decisions and the voluntary quit rate.
46 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?”
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.\footnote{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 hourly 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 −0.86. 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 RCR 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 although the bivariate results are statistically significant at the 1 percent level, the other regressions give estimates that are significant only at the 8.3, 9.5 and 12 percent levels. In these unreported regressions, the estimates for the fraction of blue collar workers are not statistically significant at any conventional level.

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

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.
1.5 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 cost(s) 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. We denote this quantity by RCE (RCEs). Alternatively, this is the product of the cost of replacing one recruit and the recruitment rate.

The relationship between RCE and establishment size is of particular interest. As mentioned above, 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, and thus these frictions are characteristic of a monopsonistic labor market. Another key question is how RCE varies with the wage. The relationship is certainly interesting in and of itself, but it also plays a key role in our later discussion.

Our strategy is to first establish in this section what sort of partial correlations exist between RCE and both the establishment size and the wage. After doing so, we interpret the relationships in light of a theoretical model.

1.5.1 Variable Definitions

In our measure of RCE, 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 RCR by the ratio of the number of recruits to the number of employees.\(^{48}\)

We also construct two alternative measures of RCE; 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 RCE 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 RCR 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 another measure of RCE. We refer to this as the imputed RCE.

Table 1.3 reports the means and standard errors of these variables and their logs.\(^{49}\) The mean value of the RCE 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

\(^{48}\)We explained above that our estimates for recruits were likely overestimates, and thus the same is likely also the case for RCEs.

\(^{49}\)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.
to the 5 year rate) to estimate recruits yields a much higher mean and a much larger
standard deviation. The mean of our imputed RCEs is reassuringly close to the mean
of our measure from interviewee’s responses. These measures are discussed in more
detail in Appendix 1.8.

1.5.2 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 1.9 displays a plot of the log of RCE 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. RCE appears to be correlated positively with establishment size, consistent with a monopsony model.50 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 RCE and establishment size.

In Figure 1.10, we plot the log of RCE 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. RCE and the wage are also strongly positively correlated.

1.5.3 Regression Analysis

Table 1.6 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 RCE 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

50Although 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.
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 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 RCE (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 RCE measure, and after including controls, the association with the log wage is much smaller in magnitude, and is no longer statistically significant.

In Table 1.5, we saw evidence of an association between RCR and the importance of seniority in promotion. As such, it is of interest to see whether controlling for this internal labor market policy affects the estimates in Table 1.6. We performed regressions analogous to those in Table 1.6 but with this added regressor, and the results are essentially the same.\footnote{We also perform regressions in which we add this control to all of the other reported tables in the paper for which it would be appropriate. Similarly, there was little effect on the results. Adding the control did have an effect on one regression mentioned in the text but not reported in table form; we discuss the details of this below.}

To investigate the key relationships of interest further, we examine individual occupational groups. The CBS asks about RCR for each occupational group. Using this information, we can construct an estimate of log RCE for each occupational group.\footnote{The CBS only has information on recruits at the establishment as a whole, so we use one fifth of the ratio of all recruits in the last five years to all employees to estimate RCE from RCR. Thus our estimates rest on the approximating assumption that the ratio of recruits to employees is constant across occupational groups.} Additionally, the CBS asks about the wage of the most recent full time hire in each of two groups: (i) non-exempt employees and (ii) professional or managerial employees. So we can use these measures to examine, by occupation, the relationships that interest us.

First consider professional and managerial workers. Table 1.7 presents results from
a regression of the log of RCE for professional and managerial workers on the log of establishment size and log of the hourly wage of the most recent full time professional or managerial hire, for several specifications. The estimates on the coefficient for the log of establishment size are a bit smaller than those in Table 1.6 and are also only very weakly statistically significant in three of the four specifications. The estimates of the coefficient on the log wage for the most recent hire in this occupation are very similar to the estimates for the wage coefficient in Table 1.6, and they are all statistically significant at the 2 percent level or lower.

We turn now to blue collar and manual labor workers. We have a measure of the log of RCE for this group, but we only have information for the log wage of the most recent nonexempt hire. We use this quantity as an approximation of the log wage of the most recent blue collar (or manual labor) hire, noting the limitation that nonexempt employees at an establishment could include clerical workers as well as some categories of sales workers. Table 1.8 displays estimates from a regression of log RCE for blue collar workers on log establishment size and the log hourly wage of the most recent nonexempt hire, for multiple specifications. The point estimates for the coefficient on the log of establishment size are all still positive, but are mostly much lower than those in Table 1.6, vary substantially, and in all but one specification are not statistically significant at any conventional level. The estimates for the coefficient of the log wage are similarly much smaller than the analogous estimates in Table 1.6 and not significant. One explanation for the decrease in the magnitude of the estimate could be classical measurement error in the log wage variable, which could arise if many of the most recent nonexempt hires are not blue collar workers but are instead from other occupational categories.

We also looked at alternative regressions by occupation, which we do not report here. The CBS has information on the hourly entry wage for workers in the most common occupational category at the establishment. We can restrict the pooled sample to those establishments for which professionals and managers are the most common occupation, and regress the log of RCE for this group on the log of the establishment size and the log hourly entry wage for the group. We consider multiple specifications, like those in Table 1.7. This gives fairly similar estimates to those for professionals above, except they are generally larger for the wage elasticity and vary more widely for the size elasticity, which is perhaps unsurprising as the sample size is reduced to about 300. We also considered analogous regressions for blue collar workers. Those size elasticity estimates are similar to the blue collar estimates in Table 1.8. The wage elasticity estimates are closer than those in Table 1.8 to the wage elasticity estimates in Table 1.6 for the full establishment, but still generally smaller than the full establishment estimates.

In summary, the results for professional and managerial workers are similar to the results in Table 1.6, with similar positive size and wage elasticities. On the other hand, estimates for blue collar workers are generally more variable and smaller for the size elasticity, and much smaller for the wage elasticity, and not generally statistically
significant. If training expenses make up a large portion of replacement costs, if there are longer training hours for professionals and managers, and if there is a correlation in pay within an establishment between new hires and those involved in training them, then these three factors together could explain the divergence, at least for the wage elasticity. Alternatively the wage disparity could be related to the greater variance in wages of professionals and managers compared to blue collar wages; if the signal-to-noise ratio is higher for the former category and if we assume again there is a correlation between wages for trainers and trainees, then the relationship between replacement costs and wages would materialize most clearly for professionals and managers.

Returning to the category of the establishment as a whole, we now 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.\footnote{This approach would not capture all effects of unobserved characteristics, however, if the relationship between each establishment’s RCE and its size changes across waves in a way other than by a simple constant additive shift in RCE across all workplaces.}

Table 1.9 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 RCE 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.\footnote{There are no 1-digit industry controls, as it is rare for an establishment to switch industry across years.} 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 RCR, and unsurprisingly it is fairly likely that an establishment will have missing data for this variable 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 1.6.

The estimate on the coefficient of bargaining coverage is negative and statistically significant at the 1 percent level. The corresponding estimate in Table 1.5 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, e.g. if the employer makes use of a union hiring hall.

We also consider regressions by occupation for the balanced panel. The regressions, which we do not report here, use the occupational log RCE and log wage measures used in the regressions by occupation above. As with the regressions in Table 1.5, sample sizes are only around 100 or smaller. The regressions for blue collar workers generally give quite variable results and the standard errors are fairly large. However, the partial correlation between log RCE for professional and managerial workers and the log wage of the most recent professional or managerial hire is positive, and statistically significant at the 6 percent level or lower. Point estimates for the partial correlation between log RCE and establishment size for the same regression are in fact negative, although not statistically significant in two out of three specifications.\textsuperscript{55}

In an unreported regression, we repeat the balanced panel regression but examine its robustness to outliers, after having examined graphs for observations likely to be exerting undue influence on the estimates. We remove only six observations with exceptional changes between survey years in the average log wage or the log RCE from the 115 used in the regression in the second column of Table 1.9. Doing so yields a large negative point estimate for the wage elasticity.\textsuperscript{56} However, for this regression, including the importance of promotion in seniority as an additional control changes the point estimate to a positive one. We conclude that the root cause of the extreme variation in estimates is very likely just the small sample size, and thus we generally do not accord as much weight to these balanced panel results as the pooled estimates.

In conclusion, we find evidence from the pooled sample that RCE is positively associated with establishment size, and the estimates are strongly statistically significant. Balanced panel regressions, however, yield point estimates that are close to zero or negative, although not statistically significant at conventional levels; however, the balanced panel regressions use fairly small samples and their resulting estimates are sensitive to outliers and choice of sets of regressors.

Regarding the wage elasticity of RCE, 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. However, as remarked above, the balanced panel estimates are fragile.

Finally, regressions for individual occupational groups show similar results for size and wage elasticities for professional or managerial workers, with the exception of negative size elasticity estimates for a balanced panel regression. But results are weaker and more variable for blue collar workers.

\textsuperscript{55}Regressions restricting to establishments with mostly professional workers use only about 30 observations, so we ignore those results.

\textsuperscript{56}Here we are removing one observation based on the value of the outcome variable. Methodological concerns should be allayed by the fact that it is clearly an unusual observation, in that the RCE changes by a factor of more than 13,000 from one wave to the next.
In the next section, we consider a formal model and discuss, in light of the model, interpretations of the partial correlations we have found for RCE.

1.6 Interpretation

In this section, we focus on a key framework for modeling RCE, due to Manning (2003). We describe the setup and then return to interpreting our results in light of the model.

1.6.1 Labor Cost Model

In the simple competitive model and the simple monopsony model, the number of workers that a firm can hire, whether it be zero workers, an infinite number or some finite number, is a function of only the wage $w$. But by, say, increasing recruitment intensity, a firm may in practice be able to hire (or retain) more workers without increasing $w$. Moreover, working harder on recruiting workers rather than increasing wages may be an especially appealing strategy for a firm if a wage increase for new workers requires a similar wage increase for current workers as well. 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.

Now consider the firm’s problem. We assume that the firm has a revenue function $Y(N)$ which incorporates the fixed price of its output. 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

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

In this setting, Manning assumes that 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, so that the firm essentially only chooses one of the two. The firm chooses $w$ and $N$ to maximize profit. Informally, one can think of this as choosing a wage and a recruiting intensity, which then translates into a choice of wage and establishment size.

$^{57}$In our discussions of the labor cost function, we draw heavily on Manning (2003) and Manning (2006).
We assume that $C(w, N)$ has continuous second partial derivatives. We also make the assumption that

$$C_w(w, N) < 0$$  \hspace{1cm} (1.2)

holds, which is certainly intuitive, as a larger wage should allow firms to work less hard to attract workers. As we have seen above, empirically there is reason to question this assumption; we discuss this issue in detail below. 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.

Returning to the profit maximization problem for the firm, consider the function $T(w, N)$ (for total labor cost) defined by

$$T(w, N) = w + C(w, N).$$

Under suitable conditions, for each fixed value $N$, there is a unique value of $w$ minimizing $T(w, N)$.\(^{58}\) Thus it is permissible to define

$$\omega(N) = \min_w (w + C(w, N)), \hspace{1cm} (1.3)$$

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

$$\pi(N) = Y(N) - \omega(N)N. \hspace{1cm} (1.4)$$

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 $N$. Observe from the assumptions of the model that choosing an $N$ is equivalent to choosing a value of output. Thus the above process can also be conceptualized as first choosing a value of output, and then choosing a wage that maximizes profit given that choice of output.

The first order condition is of the form

$$Y'(N) = \omega(N) + \omega'(N)N. \hspace{1cm} (1.5)$$

We assume that the second order condition holds.

\(^{58}\)We have $T_w(w, N) = 1 + C_w(w, N)$ and $T_{ww}(w, N) = C_{ww}(w, N) > 0$. 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 $T_w(w, N_1) = 0$. As, for each fixed value $N_1$, we have that $T_{ww}(w, N_1) > 0$ for all choices of $w$, $T(w, N_1)$ is globally convex as a function of $w$, and therefore any solution of $T_w(w, N_1)$ globally minimizes $T(w, N_1)$. (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 $N_1$.)
We now note for comparison that the firm’s profit in the basic static monopsony model is given by:

$$\pi(N) = Y(N) - w(N)N.$$  \hfill (1.6)

Observe that (1.4) is very similar in form to (1.6), the expression for profit in the simple static monopsony model, except that in (1.4), \(\omega(N)\) takes the place of the labor supply curve \(w(N)\) in (1.6).\(^59\) 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 1.3, 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

$$\omega'(N) = C_N(w, N).$$  \hfill (1.7)

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

We now mention one of the appealing features of the labor cost model: both the simple competitive and simple monopsony models can be nested within it. 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 to obtain the standard competitive model as a special case of the larger labor cost model, we set:

$$C(w, N) = \begin{cases} 0 & \text{if } w \geq w_c \\ \infty & \text{if } w < w_c \end{cases}$$  \hfill (1.8)

This choice of \(C(w, N)\) can be seen in Figure 1.11.\(^60\) Note that along the line \(w = w_c\), the labor cost function jumps.

Now consider \(C(w, N)\) in the context of the basic monopsony model. Recall the basic static monopsony model, which is depicted in Figure 1.3. 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:

$$C(w, N) = \begin{cases} 0 & \text{if } w \geq w(N) \\ \infty & \text{if } w < w(N) \end{cases}$$  \hfill (1.9)

This choice of \(C(w, N)\) can be seen in Figure 1.12. Observe that in this graph we

---

\(^{59}\)Likewise, the FOC (1.5) is similar to the first order condition in the static monopsony model, which one can readily check is given by \(Y'(N) = w(N) + w'(N)N\).

\(^{60}\)This is of course a degenerate choice of \(C(w, N)\), but one could clearly give a suitable real-valued approximation for which the resulting model would resemble the competitive model.
have labeled the labor supply curve \( w(N) \) in the \((w, N)\)-plane; along this curve, the labor cost function jumps.

### 1.6.2 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 (RCE) 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 1.5.3. In Table 1.6, we saw evidence from pooled regressions that RCEs 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. However, these results are not robust, as alternative specifications yield point estimates that are nearly zero or are negative, although these are not statistically significant at any usual levels.

In comparison, as mentioned earlier, Manning (2006) attempts to estimate the labor cost function using United Kingdom Institute for Personnel Development data on turnover costs.\textsuperscript{61} 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. Thus the data set he uses does not allow estimation of the ideal regression of log turnover costs per employee on the log establishment size and the log wage.

To address this issue, Manning makes some additional assumptions. He assumes that the turnover rate can be written as a linear function of the log wage. Using this and the FOC (1.5), 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 magnitude of the coefficient on employer size in the original, ideal regression. But he can infer the sign of the implied estimate on the original coefficient. 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

\textsuperscript{61}This is also discussed on pages 292-296 of Manning (2003).
model, namely (1.2), 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 RCE 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 increases wages? We offer some initial thoughts in the next section.

1.6.3 Discussion of 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.

**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 1.9 addressed heterogeneity by using the balanced panel to control for unobserved establishment characteristics.\(^{62}\) We found similar wage elasticity estimates, although not statistically significant ones. But the sample sizes

\(^{62}\)Although this is a strong design, it need not necessarily capture all unobserved heterogeneity, as noted above.
are quite small and, indeed, unreported regressions showed that the balanced panel estimates are fragile— for example, removal of a few outliers could lead to negative wage elasticity estimates.

In Tables 1.7 and 1.8, we focused on a more homogeneous set of workers, by looking at particular occupational groups, and looking at wages of workers at the time of hire. We found positive and statistically significant wage elasticity estimates for professional and managerial workers; on the other hand, wage elasticity estimates for blue collar workers, although positive, were generally much smaller and not significant. Although the basic model assumes for clarity that all wages within an establishment are the same, the wage at the time of hire is the particular wage that one would expect, when increased, to decrease labor costs. Since these regressions use 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 possible evidence for a negative wage elasticity of labor costs at the establishment level, but generally stronger evidence for a positive elasticity. We also find evidence that the wage relationship at the establishment level may be driven by the characteristics of professional and managerial workers at each establishment.

The CBS includes data on two quantities that are likely key components of replacement costs: the time it takes for new employees to reach full productivity, and the number of training hours to bring a new employee “up to speed.” Firms could exhibit heterogeneity in these characteristics. Naturally, these are quantities about which establishments make choices, e.g. a decision whether to hire experienced, immediately fully productive workers, or less skilled workers, or a choice about how many hours to train a new hire. So it may not be appropriate to attempt to control for heterogeneity with these variables. However, it is certainly of interest to investigate if and how these components might be driving our results.

Intuitively, one might like to know if, after taking out the effects of these components on replacement costs, there is still a remaining association with, for example, the wage. One way to do this in practice is to consider whether the residuals of a regression of RCE on these components are still positively correlated with wages, and also if these residuals are as strongly correlated as are the replacement costs themselves. By the Frisch-Waugh theorem, this is equivalent to controlling for these components in a regression, and examining whether or not the wage coefficient markedly decreases.

Table 1.10 presents the results from such an analysis. As our baseline regression for comparison, we use column 4 of Table 1.6. That column reports a regression of log RCE on the log size and log wage, multiple control variables including industry controls. Column 1 of Table 1.10 repeats the regression, but we suppress the estimates on variables other than the log size and the log wage. In the three other columns, we report regressions including the above control variables as well as, respectively, the number of months to reach full productivity, the number of new training hours (in hundreds), and both.

Time to full productivity seems to be the strongest factor, as it drives the wage
elasticity estimate down from 1.15 in column 1 to 0.83 in column 2, a drop of almost 30%; the estimate in column 3 on the other hand, of 1.13, is fairly close to the original estimate. The estimates of establishment size do decrease as well, but to a smaller extent, and seemingly equally from both quantities. The key finding from the regressions for the wage estimate is that the number of new training hours does not seem to contribute much to the correlation between RCE and the wage, while the time for workers to reach full productivity does.

Measurement Error and Other Sampling Bias

An alternate explanation for our 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 (1.1): 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 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. 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 RCR 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 1.11 reports results from regressing the log of this quantity on log establishment size and the log wage. As above, the first column uses only the log establish-

63See the data appendix for more detail for a similar argument.
64As above, we use the log of one plus the quantity rather than the log.
ment 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 0.21, 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 CBS 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. Also, a similar regression, but with the pooled sample, which we do not report, still gives positive (but weakly so) 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 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.

**Decomposition of Wage Elasticity Estimates**

Tables 1.12 and 1.13 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 1.6 into two parts: the wage elasticities of each of the log replacement costs per recruit (log RCR) 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 1.6. In these two tables, we have restricted to observations for which all variables required for Table 1.6 are not missing so that the sample does not change, in order to yield better comparisons.\(^{66}\)

\(^{65}\) Of course, one expects the wage elasticity to be lower than the values in Table 1.9 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.

\(^{66}\) Because the variables in Tables 1.12 and 1.13 are either used in Table 1.6 or are intermediate variables required in the construction of the log of replacement costs per employee, the lack of missing variables in Table 1.6 for a given observation implies there are no missing variables for that particular observation in the other two tables.
For technical reasons related to normalizing the log variables, the coefficients in Tables 1.12 and 1.13 do not sum to the corresponding coefficient in Table 1.6, 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. Nevertheless, we can still get a sense of the sign of these coefficients.

In both Tables 1.12 and 1.13, as in Table 1.6, 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 1.12 that the log RCR is positively associated with the establishment size and the wage. On the other hand, Table 1.13 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 with wage and size coming from the replacement costs per recruit, and negative associations from the recruitment rate. We can see from Table 1.6 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. Similarly the quit rate is decreasing in the establishment size. At the same time, RCRs are increasing in both the wage and the establishment size.

RCRs 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.

\[67\] 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 1.6, 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 RCE more standard over the goal of clean comparison in these three tables.

\[68\] Under equilibrium assumptions, the quit rate equals the recruitment rate. Observe also that of course the association of the quit rate and the wage is not necessarily a causal relationship.
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.

An Expanded Model

Consider another possible factor that could be playing a part in the positive wage elasticity findings. In the profit function (1.1), 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).

The exposition in section 1.6.1 reveals that the assumption that replacement costs (RCEs) are decreasing in the wage is important for the model. If these costs were increasing in the wage, then it follows immediately that the expression \( w + C(w, N) \) is increasing in the wage. Recalling (1.1), the profit function \( \pi(w, N) \) must then be decreasing in the wage, so that the optimal choice of the wage would be the minimum legal wage (or 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 positive association between RCE and the wage. This reasoning illustrates why one might expect to find a positive wage elasticity if a model were to expand the labor cost model.

We offer an expanded model of this type here. As before, let \( w \) denote the wage, \( N \) the establishment size and \( C(w, N) \) the replacement costs per employee. But now define the output

\[
Y(w, N)
\]

to depend on both \( w \) and \( N \). So a firm chooses \( (w, N) \) to maximize

\[
\pi(w, N) = Y(w, N) - (w + C(w, N))N.
\]

\textsuperscript{69}Technically, this need only hold if the profit function is \textit{strictly} decreasing, which of course follows if RCEs are strictly increasing in the wage.
The resulting first order conditions are:

\[
\frac{\partial Y}{\partial w} = N \left(1 + \frac{\partial C}{\partial w}\right), \quad \text{and}
\]

\[
\frac{\partial Y}{\partial N} - w = N \frac{\partial C}{\partial N} + C(w, N).
\]

One expects that \( \frac{\partial Y}{\partial w} > 0 \) would hold under an efficiency wage framework, as an increasing wage would increase output. Moreover, clearly one also expects \( \frac{\partial Y}{\partial N} > 0 \) to hold, as adding more workers should increase output. We observe from the first order conditions that one must have \( \frac{\partial C}{\partial w} > -1 \). Also, there is no clear condition implied by the first order conditions showing that \( \frac{\partial C}{\partial N} \) need be of one sign in particular.

Empirically, on the other hand, we have found that \( \frac{\partial C}{\partial w} > 0 \) and \( \frac{\partial C}{\partial N} > 0 \) hold, for most specifications. Unfortunately, we have very weak data on output, so we can not test this model directly, but we offer it as a possible starting point for future work.

1.7 Conclusion

In this paper, we study employee replacement costs and their relationships with multiple workplace characteristics. In the California Business Survey (CBS), 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 than for a negative association. Moreover, these associations are similarly positive
for professional and managerial workers, but weak for blue collar workers.

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; however, this may in practice only be an anomaly for professional and managerial workers. We have offered some ideas for how to reconcile the findings by expanding the model and also have offered some supporting evidence. These issues, however, clearly call for further research.
1.8 Data Appendix

1.8.1 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. Finally, there is a measure giving the wage for the most recent full time hire in each of two groups: (i) non-exempt employees and (ii) professional or managerial employees. This can be answered either as an hourly or annual wage; we convert it to an hourly wage.

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; 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.

1.8.2 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. 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. 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 1.3, 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.\(^70\) 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.

---

\(^{70}\) These correlations use the survey weights. The respective sample sizes are 808, 258 and 550.
Figure 1.1: Mean Replacement Costs (RCR)

Notes: Calculations use pooled sample of 2003 and 2008 waves of the CBS, 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.
Figure 1.2: A Simple Labor Market Model with a Fixed Wage and Increasing Replacement Costs
Figure 1.3: The Classical Static Monopsony Model

\[
\begin{align*}
N & \rightarrow \text{Wage} \\
& \quad \text{Y}'(N) = \text{MRPL} \\
& \quad w(N) \\
& \quad w_0 \\
& \quad Y'(N) = \text{MRPL} \\
& \quad N_0 \\
& \quad \text{Employees (N)}
\end{align*}
\]
Figure 1.4: Mean RCR, by Survey Wave

Notes: Costs are in 2003 dollars. Lines at the top of each bar represent 95% confidence intervals for the mean. 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 1.5: Histograms of RCR (See text for details.)
Figure 1.6: Mean and Median of RCR 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. 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 1.7: Histogram of Training Hours for New Employees, Pooled Sample
Figure 1.8: 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 1.9: Scatter Plot of Log of RCE by Log of Establishment Size
Figure 1.10: Scatter Plot of Log of RCE by Log Wage
Figure 1.11: The Labor Cost Function for the Basic Competitive Model

![Competitive Model Diagram]

Figure 1.12: The Labor Cost Function for the Basic Monopsony Model

![Monopsony Model Diagram]
Table 1.1: Estimates of Replacement Costs (RCR, Dollar Value or Percentage of Total Labor Costs)

<table>
<thead>
<tr>
<th>Estimate (%/$)</th>
<th>Description</th>
<th>Year</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.3%</td>
<td>All Employees, Single Company (International Harvester)</td>
<td>1951</td>
<td>Oi (1962)</td>
</tr>
<tr>
<td>4.1%</td>
<td>“Common Laborers”, Single Company (International Harvester)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$11,411</td>
<td>Los Angeles, Production Workers</td>
<td>1980</td>
<td>Hamermesh (1993)</td>
</tr>
<tr>
<td>$30,793</td>
<td>Los Angeles, Salaried Workers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1,723</td>
<td>Secretaries, at Large Employers</td>
<td>1979</td>
<td>Hamermesh (1993)</td>
</tr>
<tr>
<td>$5,576</td>
<td>College Graduates, at Large Employers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>150%-250%</td>
<td>Large Pharmaceutical Firm (Merck &amp; Co.)</td>
<td>1981</td>
<td>Solomon (1988)</td>
</tr>
<tr>
<td>$8,358</td>
<td>Firing Costs, France</td>
<td>1996</td>
<td>Kramarz and Michaud (2010)</td>
</tr>
<tr>
<td>$155</td>
<td>Hiring Costs, France</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$8,968</td>
<td>Managerial Workers, Large Firms, United Kingdom</td>
<td>1997–1999</td>
<td>Manning (2003)</td>
</tr>
<tr>
<td>$1,742</td>
<td>Unskilled Workers, Large Firms, United Kingdom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$8,300</td>
<td>Workers with a Vocational Degree, Switzerland</td>
<td>2002–2004</td>
<td>Blatter, Mühlemann and Schenker (2008)</td>
</tr>
<tr>
<td>$1,068</td>
<td>Average Annual Learning Expenditure per Employee</td>
<td>2008</td>
<td>American Society for Training and Development (Paradise, 2009)</td>
</tr>
<tr>
<td>2.31%</td>
<td>Annual Learning Expenditures, Percentage of Total Payroll</td>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>2.24%</td>
<td>Annual Learning Expenditures, Percentage of Total Payroll</td>
<td>2008</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Percentages refer to the percentage of total labor costs or annual wage costs or salary. See text for additional detail. Dollar figures are in 2003 dollars.
Table 1.2: Sampling Results and Response Rates in the CBS

<table>
<thead>
<tr>
<th></th>
<th>2003</th>
<th>2008 Panel Sample</th>
<th>2008 New Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampled Establishments</td>
<td>2806</td>
<td>1080</td>
<td>1072</td>
</tr>
<tr>
<td>2008 Panel Sample Still in Business in CA</td>
<td></td>
<td></td>
<td>1016</td>
</tr>
<tr>
<td>Number Meeting Eligibility Criteria</td>
<td>2200</td>
<td>868</td>
<td>849</td>
</tr>
<tr>
<td>Interview Completed</td>
<td>1080</td>
<td>358</td>
<td>294</td>
</tr>
<tr>
<td>Response Rate (Unweighted)</td>
<td>49.1%</td>
<td>41.2%</td>
<td>34.6%</td>
</tr>
</tbody>
</table>
Table 1.3: Descriptive Statistics for Key Variables, Pooled Sample

<table>
<thead>
<tr>
<th></th>
<th>Mean of Level</th>
<th>Mean of Log</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Std. Error)</td>
<td>(Std. Error)</td>
</tr>
<tr>
<td></td>
<td>[Std. Dev.]</td>
<td>[Std. Dev.]</td>
</tr>
<tr>
<td>Replacements Costs Per Recruit (RCR)</td>
<td>4039.01</td>
<td>6.41</td>
</tr>
<tr>
<td></td>
<td>(419.7)</td>
<td>(0.2)</td>
</tr>
<tr>
<td></td>
<td>[9799.8]</td>
<td>[2.7]</td>
</tr>
<tr>
<td>Establishment Size</td>
<td>43.83</td>
<td>2.57</td>
</tr>
<tr>
<td></td>
<td>(2.81)</td>
<td>(0.04)</td>
</tr>
<tr>
<td></td>
<td>[75.73]</td>
<td>[1.07]</td>
</tr>
<tr>
<td>Recruits (1 Yr Rate)</td>
<td>6.9</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(0.04)</td>
</tr>
<tr>
<td></td>
<td>[29.79]</td>
<td>[1.1]</td>
</tr>
<tr>
<td>Recruits (5 Yr Rate)</td>
<td>4.38</td>
<td>1.11</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.03)</td>
</tr>
<tr>
<td></td>
<td>[57.33]</td>
<td>[0.81]</td>
</tr>
<tr>
<td>Average Hourly Wage</td>
<td>18.55</td>
<td>2.67</td>
</tr>
<tr>
<td></td>
<td>(0.52)</td>
<td>(0.02)</td>
</tr>
<tr>
<td></td>
<td>[10.26]</td>
<td>[0.43]</td>
</tr>
<tr>
<td>Hourly Entry Wage</td>
<td>13.19</td>
<td>2.43</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.03)</td>
</tr>
<tr>
<td></td>
<td>[9.13]</td>
<td>[0.51]</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.25]</td>
<td></td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.33]</td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.27]</td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.27]</td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.24]</td>
<td></td>
</tr>
<tr>
<td>Fraction Covered by Collective Bargaining</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.15]</td>
<td></td>
</tr>
<tr>
<td>Offer Health Insurance Plan</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.38]</td>
<td></td>
</tr>
<tr>
<td>Training Department</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.4]</td>
<td></td>
</tr>
<tr>
<td>New Training Hours</td>
<td>65.59</td>
<td>3.48</td>
</tr>
<tr>
<td></td>
<td>(4.1)</td>
<td>(0.08)</td>
</tr>
<tr>
<td></td>
<td>[76.25]</td>
<td>[1.42]</td>
</tr>
<tr>
<td>Months to Full Productivity, New Professional</td>
<td>2.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[2.85]</td>
<td></td>
</tr>
<tr>
<td>Months to Full Productivity, New Non-Professional</td>
<td>2.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[2.48]</td>
<td></td>
</tr>
<tr>
<td>Importance of Seniority in Promotion</td>
<td>2.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[1.01]</td>
<td></td>
</tr>
<tr>
<td>Measures of Replacement Costs Per Employee (RCE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replacements Costs Per Employee (RCE)</td>
<td>439.17</td>
<td>4.41</td>
</tr>
<tr>
<td></td>
<td>(50.5)</td>
<td>(0.1)</td>
</tr>
<tr>
<td></td>
<td>[120.3]</td>
<td>[2.2]</td>
</tr>
<tr>
<td>Replacements Costs Per Employee (1 Yr Rate)</td>
<td>632.22</td>
<td>3.85</td>
</tr>
<tr>
<td></td>
<td>(78.72)</td>
<td>(0.18)</td>
</tr>
<tr>
<td></td>
<td>[205.44]</td>
<td>[2.92]</td>
</tr>
<tr>
<td>Imputed Replacements Costs Per Employee</td>
<td>445.24</td>
<td>4.91</td>
</tr>
<tr>
<td></td>
<td>(45.16)</td>
<td>(0.1)</td>
</tr>
<tr>
<td></td>
<td>[1117.59]</td>
<td>[1.65]</td>
</tr>
</tbody>
</table>

Notes: Mean values of variables and their logs (for some variables) in the table using a pooled sample of both waves of the CBS. See text for details.
Table 1.4: Replacement Costs Per Recruit (RCR)

<table>
<thead>
<tr>
<th></th>
<th>Mean (Std. Error)[Std. Dev.]</th>
<th>Pooled CBS</th>
<th>Cross Section</th>
<th>Balanced Panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Workers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4039</td>
<td>4529</td>
<td>3177</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(420)[9800]</td>
<td>(544)[10782]</td>
<td>(655)[7712]</td>
</tr>
<tr>
<td>Professional and</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Managerial</td>
<td></td>
<td>7051</td>
<td>7558</td>
<td>5992</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(730)[15643]</td>
<td>(950)[16902]</td>
<td>(1047)[12575]</td>
</tr>
<tr>
<td>Blue Collar</td>
<td></td>
<td>1928</td>
<td>2341</td>
<td>1246</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(475)[12290]</td>
<td>(753)[15399]</td>
<td>(165)[2853]</td>
</tr>
<tr>
<td>Clerical (2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only)</td>
<td></td>
<td>2587</td>
<td>2516</td>
<td>2516</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(269)[5320]</td>
<td>(447)[5849]</td>
<td></td>
</tr>
<tr>
<td>Sales (2003 Only)</td>
<td></td>
<td>4741</td>
<td>4976</td>
<td>4976</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(698)[16377]</td>
<td>(932)[14823]</td>
<td></td>
</tr>
<tr>
<td>Clerical and Sales</td>
<td></td>
<td>2981</td>
<td>3288</td>
<td>2406</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(269)[8530]</td>
<td>(361)[9954]</td>
<td>(380)[4786]</td>
</tr>
</tbody>
</table>

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. “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 1.5: Regression of Changes on Changes, Log of RCR on Importance of Seniority in Promotion

<table>
<thead>
<tr>
<th>Variable (Differenced)</th>
<th>Log of RCR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Std. Error)</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Importance of Seniority in Promotion</td>
<td>-1.03</td>
</tr>
<tr>
<td></td>
<td>(0.31) [0.001]</td>
</tr>
<tr>
<td>Log Establishment Size</td>
<td>0.099</td>
</tr>
<tr>
<td></td>
<td>(0.24) [0.67]</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>2.31</td>
</tr>
<tr>
<td></td>
<td>(1.38) [0.095]</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>(3.56) [0.85]</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>(2.34) [0.34]</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td>(1.35) [0.045]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>2.47</td>
</tr>
<tr>
<td></td>
<td>(3.26) [0.45]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>(2.45) [0.7]</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>-1.57</td>
</tr>
<tr>
<td></td>
<td>(2.2) [0.48]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.076</td>
</tr>
<tr>
<td></td>
<td>(0.54) [0.89]</td>
</tr>
<tr>
<td>N</td>
<td>130</td>
</tr>
</tbody>
</table>

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. Weights used are 2003 establishment weights adjusted for nonresponse.
Table 1.6: Regression of Log of RCE on Log Establishment Size and Log Wage, Pooled Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Establishment Size</td>
<td>0.37</td>
<td>0.39</td>
<td>0.36</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>(0.11) (0.001)</td>
<td>(0.12) (0.001)</td>
<td>(0.12) (0.004)</td>
<td>(0.12) (0.001)</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>1.43</td>
<td>1.12</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3)</td>
<td>(0.36) (0.002)</td>
<td>(0.36) (0.001)</td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td></td>
<td></td>
<td>−0.3</td>
<td>−0.085</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.61) (0.61)</td>
<td>(0.61) (0.89)</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td></td>
<td>0.071</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.84) (0.93)</td>
<td>(0.8) (0.43)</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>−0.54</td>
<td>−0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.54) (0.32)</td>
<td>(0.57) (0.84)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>−0.47</td>
<td>−0.031</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.8) (0.55)</td>
<td>(0.77) (0.97)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.14</td>
<td></td>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>(0.21) (0.5)</td>
<td></td>
<td>(0.2) (0.62)</td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>−0.9</td>
<td>−0.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.62) (0.15)</td>
<td>(0.57) (0.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Offer Health Insurance Plan</td>
<td>0.71</td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.34) (0.04)</td>
<td>(0.34) (0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.3</td>
<td>−0.64</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.42) (0)</td>
<td>(0.86) (0.46)</td>
<td>(1.14) (0.9)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1013</td>
<td>1000</td>
<td>879</td>
<td>878</td>
</tr>
<tr>
<td>One Digit Industry Fixed Ef-</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>fects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 one fifth of the ratio of the number 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 1.7: Regression of Log of RCE for Professionals and Managers on Log Establishment Size and Log Wage of Most Recent Professional or Managerial Hire, Pooled Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Establishment Size</td>
<td>0.39</td>
<td>0.22</td>
<td>0.23</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>(0.11)[0.001]</td>
<td>(0.14)[0.11]</td>
<td>(0.16)[0.14]</td>
<td>(0.15)[0.1]</td>
</tr>
<tr>
<td>Log Wage for Most Recent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional Hire</td>
<td>1.21</td>
<td>1.14</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.36)[0.001]</td>
<td>(0.45)[0.011]</td>
<td>(0.45)[0.016]</td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>−1.23</td>
<td>−1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.03)[0.23]</td>
<td>(1.08)[0.34]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>−0.92</td>
<td>−0.995</td>
<td>−0.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.96)[0.34]</td>
<td>(0.95)[0.92]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>−0.36</td>
<td></td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.62)[0.56]</td>
<td></td>
<td>(0.67)[0.95]</td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>−0.42</td>
<td>−0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.84)[0.62]</td>
<td>(0.82)[0.9]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>−0.056</td>
<td>−0.083</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3)[0.85]</td>
<td>(0.31)[0.79]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>−1.94</td>
<td>−1.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.83)[0.02]</td>
<td>(0.82)[0.026]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Offer Health Insurance Plan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.79</td>
<td></td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.63)[0.21]</td>
<td></td>
<td>(0.62)[0.29]</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.76</td>
<td>0.78</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.43)[0.43]</td>
<td>(0.99)[0.43]</td>
<td>(1.06)[0.31]</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1079</td>
<td>926</td>
<td>829</td>
<td>829</td>
</tr>
<tr>
<td>One Digit Industry Fixed Effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

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 one fifth of the ratio of the number of all 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 1.8: Regression of Log of RCE for Blue Collar Workers on Log Establishment Size and Log Wage of Most Recent Nonexempt Hire, Pooled Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Establishment Size</td>
<td>0.23</td>
<td>0.16</td>
<td>0.064</td>
<td>0.11</td>
</tr>
<tr>
<td>Log Wage for Most Recent Nonexempt Hire</td>
<td>0.36</td>
<td>0.18</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td></td>
<td></td>
<td>0.39</td>
<td>0.35</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td></td>
<td></td>
<td>−1.66</td>
<td>−1.42</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td></td>
<td></td>
<td>−1.55</td>
<td>−1.36</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td></td>
<td></td>
<td>−0.45</td>
<td>−0.42</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.046</td>
<td></td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td></td>
<td></td>
<td>−0.68</td>
<td>−0.93</td>
</tr>
<tr>
<td>Offer Health Insurance Plan</td>
<td></td>
<td></td>
<td>0.44</td>
<td>0.43</td>
</tr>
<tr>
<td>Constant</td>
<td>2.76</td>
<td>2.19</td>
<td>3.91</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>995</td>
<td>906</td>
<td>798</td>
<td>797</td>
</tr>
<tr>
<td>One Digit Industry Fixed Effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

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 one fifth of the ratio of the number of all 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 1.9: Regression of Changes on Changes, Log of RCE on Log Establishment Size and Log Wage

<table>
<thead>
<tr>
<th>Variable (Differenced)</th>
<th>(1) Log of RCE (Std. Error)</th>
<th>(2) Log of RCE (Std. Error)</th>
<th>(3) Log of RCE (Std. Error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Establishment Size</td>
<td>−0.083 (0.21) [0.69]</td>
<td>−0.058 (0.19) [0.76]</td>
<td>−0.27 (0.2) [0.16]</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>1.84 (1.24) [0.14]</td>
<td>1.2 (1.12) [0.29]</td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>−7.21 (2.79) [0.01]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>2.21 (1.93) [0.25]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>2.45 (1.52) [0.11]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>1.68 (2.79) [0.55]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.31 (1.84) [0.87]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>−3.05 (2.33) [0.19]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>−0.11 (0.5) [0.83]</td>
<td>−0.18 (0.46) [0.7]</td>
<td>−0.098 (0.37) [0.79]</td>
</tr>
<tr>
<td>N</td>
<td>117</td>
<td>115</td>
<td>93</td>
</tr>
</tbody>
</table>

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. Weights used are 2003 establishment weights adjusted for nonresponse.
Table 1.10: Regression of Log of RCE on Log Establishment Size and Log Wage, Pooled Sample, Effects of Time to Full Productivity and New Training Hours

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Establishment Size</td>
<td>0.39</td>
<td>0.35</td>
<td>0.36</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>(0.12)[0.001]</td>
<td>(0.13)[0.006]</td>
<td>(0.12)[0.004]</td>
<td>(0.13)[0.013]</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>1.15</td>
<td>0.83</td>
<td>1.13</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>(0.36)[0.001]</td>
<td>(0.38)[0.03]</td>
<td>(0.37)[0.002]</td>
<td>(0.38)[0.026]</td>
</tr>
<tr>
<td>Months to Full Productivity</td>
<td>0.14</td>
<td></td>
<td></td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>(0.05)[0.006]</td>
<td></td>
<td></td>
<td>(0.056)[0.006]</td>
</tr>
<tr>
<td>New Training Hours, Hundreds</td>
<td></td>
<td></td>
<td>-0.027</td>
<td>-0.085</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.21)[0.89]</td>
<td>(0.21)[0.69]</td>
</tr>
<tr>
<td>N</td>
<td>878</td>
<td>737</td>
<td>845</td>
<td>713</td>
</tr>
</tbody>
</table>

Notes: Estimates of coefficients for various regressors are suppressed; these variables are: bargaining coverage, fraction professional or managerial, fraction blue collar, fraction college degree, fraction temporary, fraction part time, offering a health insurance plan, and eight one digit industry fixed effects. 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. Replacement costs per employee are estimated using the replacement cost per recruit and multiplying by one fifth of the ratio of the number 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 1.11: Regression of Changes on Changes, Log of Replacement Costs Net of Production Value Per Employee

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
</tr>
<tr>
<td>Log Establishment Size</td>
<td>−0.14 (0.14)[0.31]</td>
</tr>
<tr>
<td></td>
<td>−0.16 (0.14)[0.26]</td>
</tr>
<tr>
<td></td>
<td>−0.34 (0.24)[0.16]</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>−0.42 (0.96)[0.66]</td>
</tr>
<tr>
<td></td>
<td>−1.26 (1.01)[0.21]</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>−10.9 (3.87)[0.005]</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.12 (2.08)[0.96]</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>0.79 (1.39)[0.57]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>1.86 (2.66)[0.48]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>−2.38 (2.22)[0.28]</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>−4.48 (2.78)[0.11]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.064 (0.4)[0.87]</td>
</tr>
<tr>
<td></td>
<td>0.088 (0.38)[0.82]</td>
</tr>
<tr>
<td></td>
<td>0.15 (0.4)[0.71]</td>
</tr>
<tr>
<td>N</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>81</td>
</tr>
</tbody>
</table>

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 costs net of production value per employee. Weights used are 2003 establishment weights adjusted for nonresponse.
Table 1.12: Regression of Log of RCR on Log Establishment Size and Log Wage

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Establishment Size</td>
<td>0.58</td>
<td>0.61</td>
<td>0.51</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.15)</td>
<td>(0.16)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>1.79</td>
<td>1.52</td>
<td>1.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(0.45)</td>
<td>(0.46)</td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>−0.11</td>
<td></td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.77)</td>
<td></td>
<td>(0.77)</td>
<td></td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.32</td>
<td>1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.1)</td>
<td>(1.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>−0.077</td>
<td></td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.72)</td>
<td></td>
<td>(0.75)</td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>−1.08</td>
<td></td>
<td>−0.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.08)</td>
<td></td>
<td>(1.04)</td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.23</td>
<td></td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td></td>
<td>(0.23)</td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>−1.42</td>
<td></td>
<td>−1.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.74)</td>
<td></td>
<td>(0.67)</td>
<td></td>
</tr>
<tr>
<td>Offer Health Insurance Plan</td>
<td>0.99</td>
<td></td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.47)</td>
<td></td>
<td>(0.45)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>4.68</td>
<td>−0.26</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.54)</td>
<td>(1.17)</td>
<td>(1.52)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1013</td>
<td>1000</td>
<td>879</td>
<td>878</td>
</tr>
</tbody>
</table>

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. 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 1.13: Regression of Log of Recruitment Rate on Log Establishment Size and Log Wage

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Establishment Size</td>
<td>−0.0088</td>
<td>−0.0092</td>
<td>−0.0073</td>
<td>−0.0064</td>
</tr>
<tr>
<td></td>
<td>(0.0029)</td>
<td>(0.0029)</td>
<td>(0.0031)</td>
<td>(0.0032)</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>−0.021</td>
<td>−0.014</td>
<td>−0.013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0073)</td>
<td>(0.0099)</td>
<td>(0.011)</td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>−0.034</td>
<td>−0.037</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>−0.037</td>
<td>−0.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>−0.024</td>
<td>−0.018</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.027</td>
<td></td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td></td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.0003</td>
<td></td>
<td>−0.0012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td></td>
<td>(0.0068)</td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>0.037</td>
<td></td>
<td>0.034</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td></td>
<td>(0.018)</td>
<td></td>
</tr>
<tr>
<td>Offer Health Insurance Plan</td>
<td>−0.013</td>
<td></td>
<td>−0.0095</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0095)</td>
<td></td>
<td>(0.0096)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.14</td>
<td>0.2</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.024)</td>
<td>(0.032)</td>
<td></td>
</tr>
</tbody>
</table>

| N                                  | 1013  | 1000  | 879   | 878   |
|                                    |       |       |       |       |
| One Digit Industry 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. 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.
Chapter 2

Wage Effects of High Performance Workplace Practices

2.1 Introduction

High Performance Workplace Organization (HPWO) practices refer to a broadly-defined set of changes in workplace practices, often involving flexible policies, that have been adopted in many firms in the U.S. in the past few decades. HPWO practices include, for example, employee meetings to discuss workplace problems, self-managed teams, and job rotation. These practices were widely adopted in the 1990s and offered the possibility of large improvements in productivity. Indeed, such improvements, suitably implemented, did seem to be generally associated with efficiency gains. (See, for example, a survey by Ichniowski et al.(Ichniowski, Kochan, Levine, Olson and Strauss 1996) and further work by Applebaum et al.(Applebaum, Bailey, Berg and Kalleberg 2001).)

A natural question arises as to how the adoption of HPWO practices affects worker wages. Multiple authors have studied this question. For example, Osterman has studied this issue in two articles(Osterman 2000, Osterman 2006). His results exhibit heterogeneity: in the first article, Osterman finds no connection between HPWO adoption and wage increases, but in the second article he finds such a connection when restricting specifically to manufacturing establishments. In general, the empirical work on this question does not provide a clear consensus about the relationship. Moreover, there have not been very many surveys addressing these issues. We aim to add to this understanding by studying the relationship between HPWO practices and wages, using a recent survey of California establishments.

It is important to emphasize that although we believe that there are clear theoretical arguments for causation going in the direction from these workplace practices to wages, our results are certainly only associations. Certainly it could be the case that the establishments which have the wherewithal to adopt HPWO practices might
be exactly those establishments that can pay high wages. Nevertheless, evidence of associations is obviously quite valuable, especially as it is likely very difficult, if not impossible, to randomly assign labor market practices to firms, and the existence of even natural experiments seems extremely unlikely.

In the next section, we discuss some of the theoretical issues involved and look at past work on this subject. In Section 2.3, we describe the survey data we use. Section 2.4 reports the key results. The final section concludes.

### 2.2 High Performance Workplace Practices: Theory and Related Literature

It is important to first note that there is no universally agreed upon definition of HPWO practices, but rather that they refer to a loosely defined set of policies. Some of these were mentioned above. Another such policy is Total Quality Management (TQM), which one survey defined for respondents as a “quality control approach that emphasizes the importance of communications, feedback, and teamwork”.\(^1\) (This last practice is also sometimes referred to as Formal Quality Management (FQM).) Other HPWO practices include changing the number of managerial levels in an organization and changing the span of control, i.e. the ratio of the number of employees to the number of managers.

In our study, we look at four practices available in our data set, namely job rotation, self-managed teams, FQM and discussion groups. In investigations of HPWO practices, often the effects of individual practices are studied, while on other occasions some index of the extent of adoption of the practices is used. In our work, we investigate both individual practices and a summary measure.

There are many ways that HPWO practices could affect wages. For the moment, we discuss wages for a production or service worker at an establishment; we consider managerial wages below.

One way HPWO policies could influence wages is that implementing the practices could involve more worker ability and thus require a higher skill workforce; this would drive up the wage establishments would need to pay to attract such a class of workers. For example, a larger skill set might be needed for job rotation, or for helping to guide a team previously led by a manager. This is essentially a human capital theory explanation. But observe that human capital theory would also suggest that wages might actually be driven down, according to the mechanism suggested by Becker (Becker 1993) in which workers bid down entry wages in order to capture future higher wages.\(^2\)

\(^1\)p.187, (Osterman 1994).

\(^2\)Many of the theoretical arguments given here are derived from lines of reasoning in (Osterman 1994), (Osterman 2000), (Osterman 2006) and (Handel and Levine 2004).
An alternative mechanism for affecting wages is the following: it could be the case that establishments invest heavily in training; they would need these employees to stay in order to get the maximal benefit from HPWO practices. In order to ensure this, in line with efficiency wage theories, establishments might increase wages as an incentive to stay with the firm.

As mentioned above, a body of research has arrived at the conclusion that HPWO practices can increase productivity. If not all of these productivity gains are passed through to the owners of the firm or to the consumer, then workers may reap some of these gains in the form of higher wages. Whether or not productivity gains pass through into wages could be dependent on the bargaining situation at the establishment; one would expect that productivity gains are more likely to be passed through if there is a union at the establishment.

Another mechanism by which HPWO practices could affect wages is also related to this issue of relative power in the workplace. It could be the case that establishments become reliant on the productivity-improving HPWO practices to retain their market position, and thus in turn are dependent on their current employees’ specialized training and knowledge in order to maintain their competitive footing. In this case, employees would have significant control and thus may be able to secure a larger percentage of productivity gains. Moreover, some HPWO practices may depend heavily on worker cooperation with the employer. Consider discussion groups in which workers suggest improvements in the production process. Employees may not be motivated to propose improvements without some benefit to themselves, such as an increased wage to engender loyalty to the firm.

On the other hand, it could be that employees feel empowerment due to discussion groups or self-managed teams, or appreciate the variety that job rotation provides, and that they are then motivated to take, or stay at, jobs with lower pay in order to take advantage of HPWO practices, in line with Rosen’s theory of compensating differentials. This mechanism would then suggest that HPWO practices could actually reduce wages.

In summary then, there are multiple theoretical arguments for a positive association between HPWO adoption and wages of production or service workers, but also a few reasons to possibly expect a negative association. Therefore, whether there is a positive or negative association is purely an empirical question. The positive influences may dominate, or the negative influences could be stronger, or perhaps the influences counterbalance each other.

We observe, as it will be relevant for the analysis below, that many of the above mechanisms could apply to either entry wages or wages for continuing workers.

So far we have discussed possible theoretical effects of HPWO practices on wages for production or service workers. Regarding managers, it could be that the use of

---

3As above, see the survey by Ichmiowski et al. (Ichmiowski et al. 1996)
4These groups are also referred to as quality circles.
HPWO practices, for example self-managed teams, leads to a reduction in the need for managerial services. If so, an establishment might require managers with fewer skills, and thus this could lead to a reduction in their wages. It is also possible that a market-wide reduction in demand for supervisors would reduce managerial wages. Alternatively, it could be the case that an establishment reduces its rolls of lower-level managers only and retains most of its higher-level supervisors; thus the only remaining managers would be higher-level, and thus the average managerial wage would actually increase.

Yet another possibility is that some workplace practices, perhaps coordinating a formal quality management program or leading discussion meetings, require increased skill on the part of the manager, thus leading to increased wages. So, as for production workers, the predicted effect on managerial wages is unclear and is a purely empirical question.

We turn to discussion of previous work on high performance workplaces. Some of the first research in the area is due to Cappelli (Cappelli 1996), who found that Total Quality Management and self-managed teams were associated with higher wages, that the number of managerial levels was associated with lower wages, although on the border of statistical significance, and finally that span of control had a small negative, but not significant, association with wages.

Key work on this topic is also due to Osterman, based on two nationwide surveys of U.S. establishments, in 1992 and 1997. Many of the initial participants were reached in the second survey, and thus a balanced panel (with 462 observations) can be formed. In 1994, Osterman (Osterman 1994) used the first wave of the data to examine the spread of, and determinants of adoption of, HPWO practices. In a follow-up study (Osterman 2000), he uses the panel, and essentially finds that the presence of HPWO practices was not associated with increased wage levels. The presence of HPWO practices was however associated with increased layoffs. Moreover, HPWO adoption generally increased from 1992 to 1997. In a 2006 paper (Osterman 2006), Osterman considers the wages of “core” blue-collar employees, in manufacturing establishments only, and finds that HPWO adoption is associated with increased wages. He also uses several instrumental variable regressions to test for robustness and obtains the same results.

Black, Lynch and Krivelyova (Black, Lynch and Krivelyova 2004) examine a panel of about 700 U.S. manufacturing establishments, and find some associations of certain types of HPWO practice adoption and higher wages for workers, though fairly few. A particularly interesting finding is an association of higher wages with the interaction of unionized status and whether or not more than 50 percent of nonmanagerial workers met regularly in discussion groups. Applebaum et al. (Applebaum et al. 2001) also give some evidence that HPWO practices are associated with higher wages. We have mentioned only a few results on the relationship between HPWO and wages: an article by Handel and Levine (Handel and Levine 2004) provides a recent survey of research on HPWO practices.
2.3 Data

2.3.1 Overview

The data set we use is a panel survey called the California Business Survey or CBS (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 CBS 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.\(^5\)

The CBS 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.\(^6\)

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 CBS 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 having at least 5 employees, and excluding government, public schools and universities, agriculture, fishing and forestry; 5

\(^5\)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.

\(^6\)The CBS 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.
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. 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. The survey designers therefore wrote an abbreviated version that omitted a number of questions from the original instrument.

Table 2.1 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+. 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 2.1 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).

Finally, we briefly discuss our use of weights with the CBS. All weights that we use account for the rate of oversampling by stratum, and account for nonresponse. 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.

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.

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.


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.

Note that percentages reported in Table 2.1 are simple unweighted percentages.
2.3.2 Key Variables and Descriptive Statistics

In Table 2.2, we give detailed descriptions for those key variables whose brief name in the regression tables might not completely convey their meanings.\textsuperscript{12} There are six central response variables, four of which are measures of the incidence of individual HPWO practices, and the fifth and sixth of which are summary variables.

The CBS asked about four workplace practices. We look at the fraction of employees involved in each of these four practices. In three of the four cases, the employee group considered is non-managerial and non-supervisory employees. The CBS asked for the fraction of non-managerial and non-supervisory employees involved in each of: (i) self-managed teams; (ii) job rotation; and (iii) regularly scheduled meetings to discuss work-related problems. The fourth HPWO practice is (iv) the fraction of all employees participating in a formal quality management (FQM) program.

The key independent variable we study is an HPWO summary measure, which we define simply as the sum of the fractions of employees involved in each of the four workplace practices. We refer to it simply as the HPWO summary measure. In some of the analyses below, we also investigate an alternative HPWO index, which we call the non-managerial HPWO summary measure. This is simply the sum of the three HPWO practices that apply only to non-managerial and non-supervisory employees. (It excludes only the FQM measure, which applies to all employees.)

Many questions in the CBS concern occupational categories. These 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. (Observe that, for example, professional and managerial workers together comprise one category.) In 2008, the clerical and sales categories were combined into one category.

We now discuss the key wage variables. The establishment average hourly wage is calculated as a mean over all employees in the establishment, using responses to questions about what fraction of employees’ wages are in particular wage categories. The average log hourly wage is computed similarly. There are several other wage measures that we use in the analysis. (All wages in the paper are hourly, and are in 2003 dollars, adjusted for inflation using the Consumer Price Index.) Two other measures are the log of the hourly entry wage and the log of the highest hourly wage in the establishment’s most common occupational category. In the analysis, we examine this for both the category of blue collar and manual labor employees, as well as the category of professional and managerial workers. The final two wage variables are the log of the starting hourly wage of the most recently hired full time non-exempt employee and the log of the starting hourly wage of the most recently hired managerial or professional employee.\textsuperscript{13} Finally, observe that the CBS does not

\textsuperscript{12}In many cases the descriptions use the original wording directly from (Survey Research Center 2003b).

\textsuperscript{13}Note that the former variable is for non-exempt employees, rather than for blue collar or manual
allow us to distinguish between wages of managerial and professional employees, as they are grouped together in all questions.

Table 2.3 reports summary statistics for key variables. The calculations are for the pooled sample of observations in the 2003 or 2008 waves of the survey, i.e. establishments present in both waves are treated as distinct observations. We present the mean and standard error for each variable, as well as its standard deviation. The calculations account for the stratified nature of the sample, using establishment weights, so that means represent estimates about the nature of the mean establishment, as opposed to the nature of the mean worker.

First consider the HPWO variables. At the mean establishment, 14.9 percent of non-managerial employees participate in self-managed teams, while 13.3 percent are involved in job rotation. 23.7 percent of all employees take part in a formal quality management program, while more than half of non-managerial employees at the mean workplace are involved in regularly scheduled meetings to discuss workplace problems. The key independent variable we study is the HPWO summary measure. Its mean is 1.07. The mean of the non-managerial HPWO summary measure, which differs from the form variable in that it is an arithmetic sum of the four measures but excluding the FQM measure, is 0.83. In summary, these workplace practices are fairly widespread.

We compare these findings to results for HPWO adoption given by Osterman for the 1997 wave of a nationwide survey, called the National Survey of Establishments (NSE) (Osterman 1994, Osterman 2000), which took place in 1992 and 1997. It differs crucially from the CBS sample we use in that it only surveys establishments with at least 50 employees, whereas the lower limit for the CBS is 5 workers. Also, the NSE is limited to for-profit establishments, whereas the CBS can include nonprofits, but the CBS does exclude the agriculture, forestry and fishing sectors and public schools and government, as mentioned above. Another key difference in the NSE survey is that many of the questions, including the questions about HPWO practices, pertain to the “core employee” group, that is, the “largest group of non-supervisory, non-managerial workers at this location who are directly involved in making the product or providing the service” at the establishment. (See p.10, (Massachusetts Institute of Technology and The Center for Survey Research 1997).) Of course the CBS is also limited to California.

To compare, we compute, for the CBS, the percentage of establishments with at least 50 employees that have strictly more than 50 percent incidence of each of these practices. Osterman presents analogous results for greater than or equal to labor employees. Also, the starting wage questions could have been answered with either an annual or monthly salaries, or with hourly wages; our measures are imputed to be estimates of hourly wages.

---

14Standard errors 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.
15The category of employees used depends on the question, as discussed above and later in this paragraph.
50 percent incidence, but we cannot do the same, because the CBS questions are asked in buckets, one of which is 25 to 50 percent, inclusive (on both ends, in fact). For teams in the CBS, we estimate the percentage as 13.7 percent, job rotation 6.7 percent, FQM 29.6 percent and discussion groups 44.6 percent. For the respective numbers in 1997, Osterman reports 38.4 percent, 55.5 percent, 57.2 percent and 57.7 percent.\textsuperscript{16} Clearly, there are large differences in some of these numbers, most notably in the measure of job rotation.

Results from past surveys have also varied significantly. The sizable difference, especially in the first comparison, could be due to many of the factors mentioned above, or due to the restriction of our data to California. One key issue is that the questions in the two surveys ask about different groups of workers. The NSE asks about the core group of workers. The CBS, on the other hand, inquires about the percentage of all employees participating in FQM, while asking about the percentage of nonmanagerial and nonsupervisory employees engaging in each of the other three HPWO practices. As the statistic Osterman chooses requires that at least 50 percent of employees participate, it is easier for a smaller category of workers to clear this bar, and thus we would expect his estimates to be higher than ours.\textsuperscript{17} Because the second comparison seems closer than the first, it does seem that the fact that the NSE focuses on core workers could be leading to the larger discrepancies in the first comparison, although clearly in both cases the job rotation measures are much smaller for the CBS. An alternate explanation of course is that there were true changes in HPWO incidence between 1997 and 2003. However, the varying settings and definitions appear the more likely explanation for a large portion of the differences.

We can examine the incidence of individual practices more closely by looking at the distributions of responses to the questions in the CBS that solicit the fraction of employees participating in each practice. Figure 2.1 presents, for each of the four HPWO practices, the frequencies of responses to each question. Frequencies are weighted by establishment weights adjusted for nonresponse. In the figure, “<25%” indicates a survey response of “less than 25%” (and clearly nonzero, as “none of them” was also an option) while “>75%” indicates a survey response of “more than 75% but less than all of them.” For all except the problem discussion meetings practice, roughly two thirds of establishments have no participating employees, with a decreasing frequency as the fractions grow larger, until a somewhat larger spike at 100 percent.\textsuperscript{18} About two fifths of establishments report that all employees participate.

\textsuperscript{16}See Table 1 of (Osterman 2000). As an attempt to bound the effect of the differing methods of reporting an exact 50 percent incidence, we also computed similar numbers for establishments with at least 50 employees having at least 25 percent incidence; this is the next lowest threshold we can use, because of the choices offered as responses in the survey. This increases the team number to 22.1 percent, job rotation to 13.2 percent, FQM to 39.7 percent and discussion groups to 58.3 percent.

\textsuperscript{17}Osterman actually makes a very similar argument in a slightly different context. See footnote 7, p. 185, (Osterman 2000).

\textsuperscript{18}We emphasize that the two thirds referred to here is a weighted number.
in problem discussion meetings, with about one fifth each reporting that none, or between 0 percent and 25 percent, participate.

Now consider some of the other variables whose means are reported in Table 2.3. The mean size of establishments in the survey is 43.8. The establishments’ largest occupational category, on average, is blue collar and manual labor workers, with 36.5 percent of the establishment’s workforce, followed by professional and managerial workers, with the clerical and sales categories trailing.

The mean of the establishment average hourly wage is around 18.5 dollars per hour. The mean establishment has only 3.1 percent of workers covered under 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.\(^{19}\) (Recall that the CBS does not include government establishments.) The number of new training hours at the mean establishment is 65.6 hours, while establishments train continuing workers 34.6 hours per year.

2.4 Results

Here we detail the regression estimates, starting with findings specifically for measures for blue collar workers. We turn afterwards to results for other classes of establishments and workers, including professionals and managers, before concluding by discussing those for the establishment as a whole.

2.4.1 Blue Collar Workers

We first investigate whether there are any relationships between wage measures for blue collar and manual labor employees and HPWO incidence at establishments with large fractions in this occupational category. As discussed above, there are many theoretical reasons to expect a positive association, but also arguments why one might see a negative correlation; it is an empirical question as to which influences predominate. To investigate this, we consider specifically the measure of the log of the hourly entry wage for workers in this category, and its relationship with the summary measure of HPWO incidence, restricted to those establishments for whom blue collar workers form their most common occupational category. We emphasize that this is the entry wage, and not an average over the occupation. We pool observations in

\(^{19}\)Our worker-weighted estimates are 12.3 percent in 2003 and 8.8 percent in 2008. This places our estimates in line with national numbers: according to the Bureau of Labor Statistics, 9.1 percent of U.S. private sector wage and salary workers in non-agricultural industries had contracts covered by a union or an employee association in 2003, and 8.5 percent in 2008. See BLS Series LU0204917300, retrieved from http://data.bls.gov/cgi-bin/srgate.
both the 2003 and 2008 waves of the survey.\footnote{As in the descriptive statistic calculations above, establishments present in both waves are treated as distinct observations. Standard errors 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.}

Table 2.4 presents results from regressions of the log entry wage at each establishment on the summary measure of HPWO incidence. The first column reports results from a simple bivariate regression, while the second includes controls for establishment characteristics, including size (in thousands), the fraction of employees covered by collective bargaining, and the fraction of full time workers eligible for health insurance benefits. Column (3) adds more controls for establishment characteristics, including the fraction of part time workers, temporary workers, and workers with a college degree; this column also includes regressors that measure the number of new and (annual) continuing training hours, both in hundreds of hours. The last two controls are included to tease apart whether HPWO practices could be responsible for higher wages, or whether increased wages might just be due to a general increase in training of which HPWO practices may just be one part. Finally, the fourth column of Table 2.4 adds one digit industry fixed effects, using the Standard Industrial Classification (SIC).

The estimates in the first two columns are both .048 and are statistically significant at the 9.7% and 12% levels, respectively. However, once we add more establishment control variables and include industry controls, the estimates are slightly larger, and are significant at the 2% level or lower. Thus, for blue collar workers, at least at the entry level, there is evidence for a positive correlation between wages and HPWO adoption.

Turning to some of the other variables, although point estimates are positive, surprisingly the establishment size is not associated with the log entry wage at any conventional level of statistical significance; after all, an employer size-wage effect has been noted in many other contexts. Bargaining coverage likewise exhibits no clear association, although generally unionization has been found to be associated with higher wages.\footnote{See (Blanchflower and Bryson 2004), for example.} On the other hand, the fraction of workers with a college degree is positively correlated with the entry wage and the fraction of part time workers is negatively correlated with it, which are both as expected.

Table 2.5 presents regressions similar to those in Table 2.4, but with a slight change in the key independent variable. The non-managerial HPWO summary measure is used here; as discussed above, it is the sum of the fractions of workers involved in each of the three HPWO practices that apply only to nonmanagerial and nonsupervisory employees. In other words, we have excluded the fraction of employees involved in FQM as the latter measure involves other employees. The regressions for each column are otherwise just like those in Table 2.4.

The estimates on this non-managerial summary measure in Table 2.5 are larger...
than the corresponding estimates in Table 2.4, but the difference is consistent with the fact that the mean of this non-managerial measure is about four fifths of the size of the other measure, as can be seen from Table 2.3. Again, as in Table 2.4, the estimates are more strongly statistically significant once controls are included. These similarities are to be expected, given the mechanical relationship between the two HPWO measures. However, the results in the table do give some evidence to back up the finding that the log entry wage for blue collar workers is associated with HPWO incidence, at least in establishments with large fractions of blue collar workers.\footnote{Moreover, we have investigated this measure as both a robustness check and as it may be a more accurate measure of the degree of nonmanagerial involvement in HPWO practices.}

We turn now to individual high performance practices. Table 2.6 presents results from regressions of the log entry wage on incidence measures for each of the four individual HPWO practices, at establishments where the largest group of workers is blue collar employees. The establishment control variables from the above regressions are included in each column, with the exception of the industry fixed effects.

Point estimates in Table 2.6 are positive for all four practices, but only the incidence of self-managed teams is positively correlated with the wage in a statistically significant fashion; the level of significance in this case is 3.9\%. Moreover, this estimate, at .16, is larger than the three others. So there is evidence that the high performance practice of having self-directed teams is associated with increases in the wage. While we do not find evidence for wage relationships with the other practices, this certainly could be only an issue of insufficient power.

As mentioned above, another wage measure reported in the CBS is the highest hourly wage in the establishment’s most common occupational category. We did a regression analysis analogous to that in Tables 2.4, 2.5 and 2.6, but with the log of this high hourly wage as the outcome variable, in place of the log entry wage. Recall that we call this the “log high wage”. For example, in the first set of regressions, the log high wage is the outcome variable, the complete HPWO summary measure is the key independent variable, and the control variables are exactly the same as in Table 2.4. We do not report the results for any of these three sets of regressions in tables. However, in summary, all of the point estimates were positive but not statistically significant at any conventional levels.

Another question in the CBS concerning wages is the starting hourly wage of the most recently hired full time non-exempt employee. We performed regressions using the log of this measure, but, although it might be an accurate measure, it is very imprecise; that is, its variance is fairly large. Under assumptions of classical measurement error, this need not lead to any bias in our estimates. However, with another form of measurement error, using this measure could lead to large biases, and the form of error is of course unknown.

Moreover, we do not have such a wage measure for blue collar workers only, but only for non-exempt workers, so the inclusion of wage rates for other workers could also
result in bias. To partially address this, one can restrict to those establishments whose most common occupational category is blue collar. However, whether restricting to this subgroup of establishments or using all establishments, regressions of the log of the starting wage for the most recent full time non-exempt hire on the HPWO summary measure yield positive, but not statistically significant, point estimates.

Naturally we cannot precisely estimate the measurement error in the log of the starting wage. However, we can get a rough sense of its size by looking at the mean of the absolute value of the difference in log starting wage for the most recent non-exempt hire between 2003 and 2008 for establishments present in both waves of the survey, and comparing to the mean of the corresponding absolute value of the difference in the log entry pay. The former mean is 0.304, while the latter is 0.204. So, from the survey data, as well as a priori from the definition, there is reason to expect that there is measurement error in the log starting wage measure.

In summary, we see that the log wage upon entry for blue collar employees is positively associated with summary measures of HPWO incidence, at establishments with a large fraction of blue collar workers. Moreover, the practice of working in self-directed teams in particular is positively correlated with higher wages for blue collar workers, and although other practices may also be similarly correlated, we do not find affirmative evidence for this.

### 2.4.2 Manufacturing Establishments and Establishments with Few Professionals

As mentioned above, Osterman (Osterman 2006) found positive correlations of the log wage with HPWO practices for a subsample of manufacturing establishments. We use the high performance summary measure to investigate if there are similar associations in the CBS sample. Table 2.7 reports regressions of the log wage on the HPWO summary measure, using manufacturing establishments in the pooled CBS data. We use controls very similar to those in Table 2.4, except that in the last two columns we now include occupational controls for the fraction of blue collar and manual labor workers and the fraction of professional and managerial employees.

In columns (1) and (2), a bivariate regression and a regression with only a few control variables yield estimates of the magnitude .04 to .05. However, once we have controlled for more establishment characteristics and for training hours for workers and industry, the estimates are larger, on the order of .075 to .08 and statistically significant at a level lower than 10 percent. For the sake of comparison, observe

---

23It might appear that it is not helpful to control for one digit industry within a manufacturing sub-sample. However, there are two one digit industries which predominantly comprise manufacturing and other SIC divisions can be viewed as manufacturing. We classify an establishment as manufacturing according to a survey response classifying the establishment as manufacturing, wholesale or retail trade or service.
that Osterman reports estimates on the order of .04 to .06, for manufacturing establishments.\textsuperscript{24} Note that the sample sizes are fairly small, on the order of 230 to 270.

There is no positive association with health insurance eligibility in the last two columns; the point estimate is in fact negative in the last column. However, we now see a statistically significant positive correlation of bargaining coverage with the log wage in the last two columns, but not in the second column. There are weak associations of both training hours measures with the wage. These last estimates could be evidence of a different relationship of these measures among manufacturing workplaces. On the other hand, the weak associations could simply be spurious results due to the small sample sizes.

It could be that in a larger manufacturing subsample or in a larger sample that is more homogeneous in some way, one might observe a positive association of wages and HPWO practices. A natural subsample to consider when investigating high performance practices is the set of establishments with not too many professional or managerial workers; as we have discussed above, there are reasons that high performance practices might affect wages differently for managers and supervisors than they do for other workers.

Table 2.8 reports results from regressions of the average log wage on the HPWO summary measure for the subsample of establishments with at most 25 percent of employees in professional or managerial occupations. The specifications are similar to those above, except that we do not control for the fraction of workers in professional or blue collar occupations.

Estimates of the coefficient on the measure of HPWO incidence are on the order of .05 to .06, and are statistically significant at the 5.7 percent level or lower. These estimates are on the high side of the range that Osterman obtains for manufacturing establishments, and in the middle of the range that we observe in Table 2.7 for manufacturing workplaces. It may be that the effects are more clearly seen here because of the larger sample we can investigate by turning from manufacturing establishments to a larger, and perhaps similar, set of workplaces.

We see positive associations of the wage with establishment size, insurance eligibility and the fraction of workers with a college degree, although the size result is not robust. As before, there is a negative association with the fraction of part time workers. Notably, there is a weak positive association with training hours for new workers, as in the manufacturing subsample. But unlike that subsample, there is no clear relationship between the wage and continuing training hours.

In Table 2.9, we restrict to the same sample of establishments, but look at re-

\textsuperscript{24}See Tables 3 and 4 of (Osterman 2006) for his estimates. It is not clear what the final sample size is for manufacturing establishments in these regressions, but due to missing variables and before restricting to manufacturing, it is 492. Perhaps restricting to manufacturing in the NSE retains more observations than we have due to the NSE consisting of establishments with 50 or more employees whereas the CBS sample need have only 5 or more employees.
gressions of the wage on individual HPWO practices. Control variables are the same as those in Table 2.8. The estimate for the coefficient on the fraction involved in self-managed teams is .17, much larger than above estimates of the summary measure, and is significant at the 8.1 percent level. The corresponding estimate for job rotation is even larger, and is significant at the 2.4 percent level. The estimates for the other two practices are .03 and .04, closer to the low range of estimates for the summary measure. Recall that the incidence of teams and job rotation is much lower than the incidence of the other two HPWO practices; this explains how the estimates for teams and job rotation here can be much different from the estimate in Table 2.8 for the summary measure.

For the team and job rotation practices, we investigate further using the balanced panel feature of the CBS. There are 311 establishments that were asked questions about wages and high performance workplace practices in both years of the survey. Using this information, we can regress the changes in the log wage from 2003 to 2008 on changes in the HPWO summary measure, which is equivalent to controlling for fixed effects of individual establishments. This allows one to control for unobserved characteristics of the establishment to some extent. We also include a constant, to allow for a change in the log wage that is uniform across establishments, as well as changes in multiple control variables.

Table 2.10 reports regressions of the change in the average log wage from 2003 to 2008 on changes in these measures. All of the regressions restrict to those establishments with at most 25 percent of employees professional or managerial in 2003. Columns (1) and (2) concern self-managed teams. Column (1) is the simple bivariate regression of the change in wage on the change in the incidence of teams and a constant. The second column has the same form, but now includes various control variables; to be clear, the regressors are changes in the values of the corresponding control variable. Columns (3) and (4) are analogous, but with the change in the incidence of job rotation as the key regressor rather than the change in the incidence of teams.

In column (1), the magnitude of the estimate is fairly large, with a value of .19. The estimate remains largely the same after inclusion of controls. In both columns (1) and (2), the estimate is statistically significant at a level of 2.2 percent or lower. The sample size is very small however, especially after first using the balanced panel, and then further restricting to this particular subsample of workplaces.

On the other hand, the point estimates are actually negative (and not statistically significant) for job rotation, unlike the positive estimates in Table 2.9. Observe that values of about .08 or higher for the coefficient on job rotation can be ruled out, with 95 percent confidence.

---

25 Naturally, a smaller number responded to all relevant questions in both waves.
2.4.3 Professionals and Managers

We turn now to relationships for professional and managerial workers. Table 2.11 presents results for regressions of the log of the hourly entry wage for professional and managerial workers on the HPWO summary measure, restricted to those establishments for which they comprise the most common occupational category. The first column is a bivariate regression, while columns (2) and (3) add control variables as in above tables. The final column includes one digit industry fixed effects as well.

We see strong positive correlations between the wage and HPWO incidence in the first two columns in Table 2.11. However, in the last two specifications, the point estimate on the summary measure is positive but only very weakly statistically significant. So here there is not particularly strong evidence for a positive relationship between HPWO adoption and wages. However, the magnitude of the estimates are around .075 to .092 and so are larger than many of the estimates above. As the sample sizes are fairly small (around 300 to 350), it could be that a larger sample would have the power to detect an association. Regarding the control variables, there are strong positive relationships between the wage and both health insurance eligibility and the fraction of workers with a college degree, and a strong negative relationship between the wage and the fraction of part time workers.

Now consider the individual high performance practices. Table 2.12 presents regressions of the log hourly entry wage for professionals and managers on each of the four measures of incidence of the individual practices, again restricted to workplaces where managers and professionals comprise the most common worker group.

Interestingly, we see some weak evidence, statistically significant at the 11 percent level, that the fraction of nonmanagerial and nonsupervisory workers in self-directed teams is negatively related to wages for managers and professionals, at least at the entry level. This is consistent with the mechanism mentioned above in which the introduction of self-managed teams for nonsupervisory workers reduces the need for managers and thus lowers their wages. Of course we should bear in mind that, as remarked above, we only have data for the commingled group of professionals and managers. Also, we see a very strong positive relationship between wages for this group and the fraction of all employees involved in formal quality management, which is statistically significant at the 0.3% level. The magnitude of the correlation is large relative to the other estimates we have seen above; to get a sense of the size, a theoretical change in an establishment from none to all employees participating in FQM would be associated with a 31 percent increase in professional and managerial wages.26 Finally, there is a weak positive association between increased professional/managerial wages and the fraction of nonsupervisory employees participating in meetings to discuss work-related problems.

We turn now to the log of the highest hourly wage for professionals and managers,

\[ \text{This is } \exp(0.27); \text{ to use the estimate, one must of course use assume the regression provides a good estimate of changes in the wage for a broad range of changes in the independent variable.} \]
which we again refer to as the log high wage. Table 2.13 presents results from regressions of the log high wage for this group on the HPWO summary measure and control variables, restricted to the set of establishments where this group is the most common occupational category. The control variables are just as in Table 2.11; the first column is a bivariate regression, the next includes a few more regressors, and the third column still more, while the final column adds industry fixed effects. Just as in Table 2.11, we see strong positive correlations between HPWO incidence and the highest wage for professionals and managers, for the bivariate regression and the specification with a few additional controls. But in columns (3) and (4), these correlations are positive but only very weakly statistically significant. As with Table 2.11, it seems quite possible that our results could be an indication that a larger sample would confirm the positive relationship, as the analysis lacks statistical power: indeed the sample sizes are only about 300. However, having seen the results of Table 2.12, the weakness of the relationship seems in large part due to the heterogeneity of the associations between the individual practices and the wage. Table 2.12 suggests that the incidence of FQM and problem discussion meetings is positively related to the wage, while that for teams is negatively related.

Table 2.14 considers the relationships of the log high wage for professionals and managers with the incidence of each individual high performance practice. Each column presents the results of a regression of the log high wage on the incidence of one such practice and several control variables, restricted to establishments whose most common occupational category is professional or managerial.

The results are fairly similar to those for the log entry wage in Table 2.12. The fraction of all employees participating in FQM is still strongly positively associated with increased wages for professionals and managers, and at very much the same level of magnitude as in Table 2.12. There is, also as in Table 2.12, no evidence of a correlation between job rotation incidence and wages and weak evidence for a positive relationship between the fraction of nonsupervisory employees participating in discussion groups and wages. One key difference is that the negative relationship between the fraction of nonsupervisory workers in self-managed teams and the wage is much stronger in Table 2.14 than in Table 2.12; here the result is statistically significant at the 4.7% level. So there is more evidence for heterogeneity in the relationships between wages and the individual HPWO practices.

Although unreported, we also considered regressions of the log of the starting wage for the most recently hired full time professional or manager on the HPWO summary measure and on the individual measures (with all establishments in the sample). For the summary measure of all HPWO practices, point estimates are actually negative, although not statistically significant; this is perhaps not surprising in view of our concerns about measurement error. Estimates for the individual practices are not statistically significant, although there is evidence for a negative relationship with the incidence of self-managed teams: the estimate is $-0.32$, which is statistically significant at the 2% level, and is fairly similar to the estimate of $-0.24$ obtained in
Table 2.14.

Summarizing our results for professionals and managers, we see weak evidence for a positive relationship between overall HPWO incidence and both the log entry and log high wage for these workers. This weakness is likely a result of the heterogeneous relationships between the individual HPWO practices and these wages, given the regression results we have seen for these practices. We find a positive correlation between the fraction of all employees participating in FQM and entry and high wages for professionals and managers, and a negative relationship with the portion of nonsupervisory workers in self-managed teams.

2.4.4 Establishment-Wide Results

We now look at the relationship between average wages at the establishment level and our high performance workplace summary measure, for the establishment as a whole, instead of for specific subgroups of workers or for specific types of establishments. Although there are theoretical reasons to expect such relationships for subgroups, we have offered no clear rationale to expect any such association at the establishment-wide level. Indeed, correlations at the level of subgroups could in fact cancel each other out, leading to no association at an aggregate level. However, the aggregate relationship is nevertheless of interest and worth investigating.

Table 2.15 presents results from regressions of the average log wage at each establishment on the summary measure of HPWO incidence. We pool observations in both the 2003 and 2008 waves of the survey.

The specifications are similar to those above. The first column reports results from a simple bivariate regression, whereas the second includes controls for establishment characteristics. Column (3) adds more controls for establishment characteristics. The final column of Table 2.15 adds one digit industry fixed effects.

The point estimates of the coefficient of the HPWO measure are positive, but not statistically significant at any conventional level. The estimates are on the order of .03 to .04. Establishment size is positively correlated with the wage. Similarly, the fraction of full time workers eligible for health insurance is unsurprisingly positively correlated with the wage. Bargaining coverage exhibits no clear association. New training hours also show no clear relationship with the wage, but the estimates for continuing training hours indicate a negative relationship with the log wage, and so does the fraction of part time workers. But to sum up regarding our main question of interest, we do not see a statistically significant positive association of the log wage with high performance practices, at the establishment as a whole, at least from the pooled sample estimates in Table 2.15.

The negative relationship found above between the incidence of teams and professional wages could be a key reason explaining, on the one hand, our finding of a positive correlation of the average log wage and HPWO adoption for establishments with few professionals or managers, and, on the other hand, our lack of finding of a
correlation between the average log wage and HPWO adoption in the establishment as a whole; it seems likely that in establishments with few professionals or managers, the negative association between team adoption and professional wages may not be an important factor in the relationship between HPWO adoption and the average establishment-wide log wage. We may be seeing evidence of heterogeneous relationships that to some extent cancel each other when we look at broader measures.\footnote{Observe that both the relationship between the HPWO summary measure and wages at establishments with few professionals, and its relationship with professional and managerial wages, are at least weakly positive. This situation may not at first appear to be capable of leading to the (small) cancellation effect needed to yield a smaller magnitude relationship between HPWO adoption and wages over all establishments. However, a complicated interplay of the heterogeneous relationships between the individual workplace practices and wages for different occupational groups is a potential explanation. It could also be explained by other potential special properties of establishments that have fairly few managers or professional workers.}

Table 2.16 reports results using the balanced panel aspect of the CBS. We regress the changes in the log wage from 2003 to 2008 on changes in the HPWO summary measure.

In column (1), we see that the basic estimate of the coefficient on the high performance measure is positive, but is far from statistically significant at any conventional level. The sample size (252) is fairly small, so it could be that a larger sample would yield a positive and statistically significant estimate. One way to delve further is to add other controls. Thus we include controls for changes in establishment size, changes in bargaining coverage and changes in full time health insurance eligibility in column (2). The third column includes controls for changes in training hours, new and continuing, while the final column includes controls for changes in other regressors.

Looking at the results in these last three columns, we see that the point estimate for the high performance measure is even negative in the last two. The sample size in the last column is 177, and the standard errors are unsurprisingly large, and so one cannot rule out even a true partial correlation of .04 (at a level of 5 percent). But we certainly find no evidence for a positive association of wages and HPWO practices.

Other estimates worth noting are the positive association of establishment size and the wage, and the negative association of the fraction of part time workers and continuing training hours with the wage. The first two results are as expected, and are evident in Table 2.15 as well. But the association with continuing training hours is not clear from theoretical predictions. Perhaps one might expect the opposite correlation in that more difficult jobs would require continual updating of skills, but it could be the case that lower-skill workers need continuing training more often. Also, the association is not seen in Table 2.15 for the pooled sample, so it could be an artifact of the small sample size, although the results are strongly statistically significant, but perhaps more likely the unobserved fixed effects are important in detecting this partial correlation.

We now consider the relationship of the individual practices with the log wage,
rather than the summary measure. Table 2.17 considers regressions of the average log wage on each HPWO practice separately, with controls as in Table 2.15, using the pooled sample. The point estimates on each practice are positive, but none of them are statistically significant at any conventional level. The estimates for teams and job rotation are larger in magnitude than the other two. There are, as in Table 2.15, and as expected, positive associations of the log wage with establishment size, full time employee health insurance eligibility, the fraction of workers with a college degree and the fraction of professional and managerial workers, and negative associations with the fraction of part-time and blue collar workers. These associations are strongly statistically significant for the most part, and all are statistically significant at the level of 6.7 percent or lower.

In summary, for the establishment as a whole, we find no evidence of any correlation between HPWO adoption and wages. So HPWO practices may be important only for specific subgroups of workers, and as we have remarked, heterogeneity could be leading to dampened findings for aggregate measures and categories.

2.4.5 Union Density

Recall the work of Black, Lynch and Krivelyova (Black et al. 2004) mentioned above, who find a positive association of wages with the interaction of union status and HPWO practice adoption. To consider this question for the CBS, we restrict to establishments (in the pooled sample) which have at least some workers covered under a collective bargaining agreement. This is actually a fairly small sample, consisting of only 284 observations. We again use the HPWO summary measure rather than measures for individual workplace practices; the regressions are performed for the establishment as a whole, rather than for a particular subclass of workers; sample sizes would be much smaller if we were to investigate subclasses.

Table 2.18 reports the results from these regressions. In the bivariate regression in column (1) and in the regression in column (2), with few controls, the estimates are large and statistically significant at the 5 percent level or lower. However, adding numerous controls in column (3) and one digit industry fixed effects in column (4), the point estimates are smaller, at .08 and .05 and no longer statistically significant at any conventional level. Overall, we see some suggestive evidence that bargaining coverage in conjunction with HPWO practices might lead to higher wages, and also that this relationship might be stronger than that between HPWO practices and wages for manufacturing establishments. But our sample of establishments with some bargaining coverage is too small to reach definitive conclusions.

28 One digit industry fixed effects however are not included in any of the regressions in Table 2.17.

29 Using the balanced panel in an unreported calculation, we do find a very weak positive correlation between changes in the average log wage and changes in the fraction of self-managed teams, but this is statistically significant at only the 20% level.
2.5 Conclusion

Using a panel survey of California businesses, the CBS, we find heterogeneity in the relationship between wages and the adoption of HPWO practices: the analysis yields evidence for a positive association between HPWO adoption and wages for particular subgroups of workers, and for particular HPWO practices.

We find reasonably strong evidence for a positive correlation between the log hourly entry wage for blue collar and manual labor workers and HPWO adoption, by examining these partial correlations for establishments at which blue collar workers comprise the largest occupational group. We tested this for robustness with an HPWO summary measure for the three practices measured for non-managerial employees. Moreover, among the individual high performance practices, the incidence of self-managed teams among non-managerial and non-supervisory workers is strongly positively correlated with increased hourly entry wages for blue collar workers.

We also examine another special category for which previous research indicates HPWO adoption might be associated with increased wages: manufacturing establishments. For such workplaces, we find only weak indications that HPWO practices correspond with increased establishment-wide measures of log wages, although the lack of a strong finding could be due to small sample size. A possibly similar class of establishments is the set of workplaces with a small fraction of professional and managerial workers. For these establishments, we find statistically significant (at the 5.7% level or lower) partial correlations of HPWO intensity and higher average log wages at the establishment. In these two types of regressions, point estimates are on the same order as, or larger than, the estimates found by Osterman (Osterman 2006) in his study of manufacturing establishments.

Within the sample of establishments with relatively few professionals and managers, we also see some weak evidence for a positive association of wages and the incidence of self-managed teams, and strong evidence for a positive association of wages and the prevalence of job rotation. However, when we turn to using the balanced panel, so that one can control for establishment fixed effects, the relationship with job rotation appears negative while that for self-managed teams seems quite strongly positive, although any interpretations should heed the small sample sizes.30

We also consider wages for professionals and managers. As noted above, we cannot separate the wages of these two types of workers using the CBS data set. There is weak evidence for a positive relationship between overall HPWO incidence and both the log entry and log high wage for these workers. This weakness appears to be explainable by heterogeneity in the relationships of wages and individual high performance practices: there is a positive correlation between the fraction of all employees participating in FQM and both entry and high wages for professionals and managers, and a negative relationship with the proportion of nonsupervisory workers in self-managed teams.

30To be clear, the sample is here still restricted to establishments with a small fraction of professional or managerial workers.
To investigate whether unions could play a role in workers’ capturing productivity gains, we also look at the general HPWO summary measure, while restricting to establishments with at least some collective bargaining coverage. Although our sample of such workplaces is fairly small, we do find suggestive evidence for such a correlation, in agreement with previous work.

Finally, we find no clear evidence of a positive association overall between HPWO adoption and log wages for the establishment as a whole; as mentioned, we have of course offered no rationale to expect any such relationship.

In summary, we observe strong indications that there is substantial heterogeneity in the relationship between HPWO practice adoption and wages. HPWO may lead to higher wages only in certain industries, such as manufacturing. Or perhaps higher wages only occur for certain types of workers, perhaps particularly for blue collar or nonmanagerial workers, but maybe also for professional workers and managers, with perhaps no increases for clerical workers. Alternatively, bargaining power may be necessary to capture these wages. Moreover, perhaps certain HPWO practices are related to wages in different ways than other practices, and perhaps even differently for separate groups of workers. For example, the incidence of self-managed teams appears to be correlated with much higher wages than other practices for blue collar workers; for professional or managerial workers on the other hand, it is associated with lower wages while participation in a formal quality management program is correlated with higher pay. Additionally, this heterogeneity may manifest in such a way that correlations at finer levels cancel each other, leading to an absence of a relationship at aggregate levels.

We remind the reader that our research design cannot establish causality, although there are numerous theoretical reasons to believe that HPWO adoption might lead to higher wages. Our evidence most strongly indicates that the relationship between high performance workplace practices and wages is extremely complex, and simple correlations of the two without taking into account other varying quantities may miss these intricacies. The results we have presented suggest that there are some environments in which HPWO adoption and wages are strongly related, and that more work is needed in order to uncover the exact conditions and mechanisms.
Figure 2.1: Frequencies of Responses for Individual HPWO Practices

Notes: Weighted frequencies are plotted by fraction of employees participating in each HPWO practice. (For FQM, it is the fraction of all employees. For all other practices, it is the fraction of non-managerial and non-supervisory employees.) “<25%” indicates a survey response of “Less than 25%” (and presumably nonzero as “none of them” was also an option) while “>75%” indicates a survey response of “more than 75% but less than all of them”. The number N indicates the number of observations in the data set that have nonmissing data for that variable. These are not histograms: although the widths are equal in the figures for ease of reading, areas do not have any conventional interpretations. Pools all observations in 2003 and 2008 waves; establishments in both waves are treated as distinct observations. Uses establishment weights adjusted for nonresponse.
Table 2.1: Sampling Results and Response Rates in the CBS

<table>
<thead>
<tr>
<th></th>
<th>2003</th>
<th>2008 Panel Sample</th>
<th>2008 New Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampled Establishments</td>
<td>2806</td>
<td>1080</td>
<td>1072</td>
</tr>
<tr>
<td>2008 Panel Sample Still in Business in California</td>
<td></td>
<td>1016</td>
<td></td>
</tr>
<tr>
<td>Number Meeting Eligibility Criteria</td>
<td>2200</td>
<td>868</td>
<td>849</td>
</tr>
<tr>
<td>Interview Completed</td>
<td>1080</td>
<td>358</td>
<td>294</td>
</tr>
<tr>
<td>Response Rate (Unweighted)</td>
<td>49.1%</td>
<td>41.2%</td>
<td>34.6%</td>
</tr>
</tbody>
</table>
Table 2.2: Definitions of Key Variables

<table>
<thead>
<tr>
<th>High Performance Workplace Practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction in Self-Managed Teams</td>
</tr>
<tr>
<td>Fraction in Job Rotation</td>
</tr>
<tr>
<td>Fraction Participating in FQM</td>
</tr>
<tr>
<td>Fraction in Problem Discussion Meetings</td>
</tr>
<tr>
<td>HPWO Summary Measure</td>
</tr>
<tr>
<td>HPWO Non-Managerial Summary Measure</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>General Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction Temporary</td>
</tr>
<tr>
<td>Fraction Part-Time</td>
</tr>
<tr>
<td>Fraction College Degree</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Labor Market Policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction under Collective Bargaining</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
</tr>
<tr>
<td>Continuing Training Hours per Year</td>
</tr>
<tr>
<td>New Training Hours</td>
</tr>
</tbody>
</table>
Table 2.3: Descriptive Statistics for Key Variables (Pooled Sample)

<table>
<thead>
<tr>
<th>Mean</th>
<th>(Std. Error) [Std. Dev.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction in Self-Managed Teams</td>
<td>0.149</td>
</tr>
<tr>
<td>Fraction in Job Rotation</td>
<td>0.133</td>
</tr>
<tr>
<td>Fraction Participating in FQM</td>
<td>0.237</td>
</tr>
<tr>
<td>Fraction in Problem Discussion Meetings</td>
<td>0.548</td>
</tr>
<tr>
<td>HPWO Summary Measure</td>
<td>1.066</td>
</tr>
<tr>
<td>Non-Managerial HPWO Summary Measure</td>
<td>0.83</td>
</tr>
<tr>
<td>Establishment Size</td>
<td>43.83</td>
</tr>
<tr>
<td>Average Hourly Wage</td>
<td>18.546</td>
</tr>
<tr>
<td>Average Log Hourly Wage</td>
<td>2.67</td>
</tr>
<tr>
<td>Log Starting Wage for Most Recent Full Time Professional or Managerial Hire</td>
<td>2.919</td>
</tr>
<tr>
<td>Log Starting Wage for Most Recent Full Time Nonexempt Hire</td>
<td>2.436</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.306</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.053</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>0.178</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.282</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>0.365</td>
</tr>
<tr>
<td>Fraction Covered by Collective Bargaining</td>
<td>0.031</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>0.758</td>
</tr>
<tr>
<td>New Training Hours</td>
<td>65.591</td>
</tr>
<tr>
<td>Continuing Training Hours per Year</td>
<td>34.648</td>
</tr>
</tbody>
</table>

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. Standard deviations are in brackets. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.4: Regression of Log Hourly Entry Wage for Blue Collar Employees on Summary HPWO Measure, Pooled Sample, Only Establishments Whose Most Common Occupation is Blue Collar

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPWO Summary Measure</td>
<td>0.048</td>
<td>0.048</td>
<td>0.053</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>(0.029)[0.097]</td>
<td>(0.031)[0.12]</td>
<td>(0.023)[0.021]</td>
<td>(0.016)[0.002]</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.06</td>
<td>0.064</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.075)[0.42]</td>
<td>(0.07)[0.36]</td>
<td>(0.062)[0.1]</td>
<td>(0.092)[0.23]</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.17</td>
<td>0.18</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.13)[0.2]</td>
<td>(0.11)[0.12]</td>
<td>(0.092)[0.23]</td>
<td>(0.092)[0.23]</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health</td>
<td>0.041</td>
<td>0.045</td>
<td>0.065</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.066)[0.54]</td>
<td>(0.057)[0.44]</td>
<td>(0.042)[0.12]</td>
<td>(0.042)[0.12]</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>−0.32</td>
<td></td>
<td></td>
<td>−0.19</td>
</tr>
<tr>
<td></td>
<td>(0.066)[0]</td>
<td>(0.059)[0.001]</td>
<td>(0.059)[0.001]</td>
<td>(0.059)[0.001]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>−0.056</td>
<td>−0.081</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.087)[0.52]</td>
<td>(0.054)[0.13]</td>
<td>(0.054)[0.13]</td>
<td>(0.054)[0.13]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.25</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.14)[0.83]</td>
<td>(0.096)[0]</td>
<td>(0.096)[0]</td>
<td>(0.096)[0]</td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.028</td>
<td>0.043</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.034)[0.42]</td>
<td>(0.03)[0.15]</td>
<td>(0.03)[0.15]</td>
<td>(0.03)[0.15]</td>
</tr>
<tr>
<td>Continuing Training Hours per Year,</td>
<td>0.022</td>
<td>−0.033</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in Hundreds</td>
<td>(0.033)[0.51]</td>
<td>(0.021)[0.12]</td>
<td>(0.021)[0.12]</td>
<td>(0.021)[0.12]</td>
</tr>
<tr>
<td>Constant</td>
<td>2.13</td>
<td>2.09</td>
<td>2.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.036)[0]</td>
<td>(0.054)[0]</td>
<td>(0.071)[0]</td>
<td>(0.071)[0]</td>
</tr>
</tbody>
</table>

N 717 713 630 630

Notes: Pools all establishments in 2003 and 2008 waves for which the most common occupational category is blue collar and manual labor; 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.5: Regression of Log Hourly Entry Wage for Blue Collar Employees on Summary Non-Managerial HPWO Measure, Pooled Sample, Only Establishments Whose Most Common Occupation is Blue Collar

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Hourly Entry Wage (Std. Error)</td>
<td>p-value</td>
<td>(p-value)</td>
<td>p-value</td>
<td>(p-value)</td>
</tr>
<tr>
<td>HPWO Nonmanagerial Summary Measure</td>
<td>0.068</td>
<td>0.071</td>
<td>0.075</td>
<td>0.067</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.081</td>
<td>0.092</td>
<td>0.11</td>
<td>0.071</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.17</td>
<td>0.18</td>
<td>0.11</td>
<td>0.22</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>0.053</td>
<td>0.056</td>
<td>0.076</td>
<td>0.061</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>-0.3</td>
<td>-0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>-0.076</td>
<td>-0.093</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.25</td>
<td>0.35</td>
<td>0.077</td>
<td>0.090</td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.029</td>
<td>0.044</td>
<td>0.077</td>
<td>0.14</td>
</tr>
<tr>
<td>Continuing Training Hours per Year,</td>
<td>0.018</td>
<td>-0.033</td>
<td>0.023</td>
<td>0.14</td>
</tr>
<tr>
<td>in Hundreds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.13</td>
<td>2.08</td>
<td>2.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.054)</td>
<td>(0.066)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>728</td>
<td>724</td>
<td>637</td>
<td>637</td>
</tr>
</tbody>
</table>

Notes: Pools all establishments in 2003 and 2008 waves for which the most common occupational category is blue collar and manual labor; 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.6: Regression of Log Hourly Entry Wage for Blue Collar Employees on Individual HPWO Measures, Pooled Sample, Only Establishments Whose Most Common Occupation is Blue Collar

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction in Self-Managed Teams</td>
<td>0.16</td>
<td>(0.076)[0.039]</td>
<td>0.096</td>
<td>(0.088)[0.27]</td>
</tr>
<tr>
<td>Fraction in Job Rotation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Participating in FQM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction in Problem Discussion Meetings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.076</td>
<td>(0.074)[0.31]</td>
<td>0.055</td>
<td>(0.074)[0.46]</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.16</td>
<td>(0.12)[0.18]</td>
<td>0.18</td>
<td>(0.12)[0.15]</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health</td>
<td>0.072</td>
<td>(0.05)[0.16]</td>
<td>0.067</td>
<td>(0.052)[0.2]</td>
</tr>
<tr>
<td>Insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>-0.3</td>
<td>(0.063)[0.9]</td>
<td>-0.3</td>
<td>(0.064)[0.9]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>-0.08</td>
<td>(0.091)[0.38]</td>
<td>-0.064</td>
<td>(0.088)[0.47]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.24</td>
<td>(0.14)[0.78]</td>
<td>0.3</td>
<td>(0.15)[0.04]</td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.028</td>
<td>(0.037)[0.44]</td>
<td>0.034</td>
<td>(0.036)[0.34]</td>
</tr>
<tr>
<td>Continuing Training Hours per Year, in Hundreds</td>
<td>0.026</td>
<td>(0.031)[0.4]</td>
<td>0.027</td>
<td>(0.038)[0.47]</td>
</tr>
<tr>
<td>Constant</td>
<td>2.08</td>
<td>(0.067)[0.9]</td>
<td>2.07</td>
<td>(0.065)[0.9]</td>
</tr>
</tbody>
</table>

N 641 642 638 643

Notes: Pools all establishments in 2003 and 2008 waves for which the most common occupational category is blue collar and manual labor; 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.7: Regression of Average Log Wage on HPWO Summary Measure, Pooled Sample, Manufacturing Establishments Only

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPWO Summary Measure</td>
<td>0.048</td>
<td>0.04</td>
<td>0.075</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.049)</td>
<td>(0.044)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.14</td>
<td>-0.072</td>
<td>-0.082</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.033)</td>
<td>(0.034)</td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.078</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health</td>
<td>0.041</td>
<td>-0.018</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.097)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td>-0.43</td>
<td>-0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>0.32</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.78</td>
<td>0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.32</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>-0.31</td>
<td>-0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.044</td>
<td>0.045</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.028)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuing Training Hours per Year, in Hundreds</td>
<td>0.057</td>
<td>0.057</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.71</td>
<td>2.43</td>
<td>2.51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.087)</td>
<td>(0.17)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>274</td>
<td>273</td>
<td>232</td>
<td>232</td>
</tr>
<tr>
<td>One Digit Industry Fixed Effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Notes: Pools all observations in 2003 and 2008 waves for which respondents state that their industry is manufacturing; 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.8: Regression of Average Log Wage on HPWO Summary Measure, Pooled Sample, Only Establishments with At Most 25 Percent Professional or Managerial Employees

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPWO Summary Measure</td>
<td>0.059</td>
<td>0.062</td>
<td>0.058</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(0.029)[0.046]</td>
<td>(0.033)[0.057]</td>
<td>(0.024)[0.016]</td>
<td>(0.022)[0.027]</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.092</td>
<td>0.18</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.074)[0.21]</td>
<td>(0.084)[0.028]</td>
<td>(0.077)[0.011]</td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.062</td>
<td>0.074</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.14)[0.65]</td>
<td>(0.14)[0.59]</td>
<td>(0.12)[0.89]</td>
<td></td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>0.27</td>
<td>0.15</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.069)[0]</td>
<td>(0.067)[0.028]</td>
<td>(0.055)[0.003]</td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>–0.56</td>
<td>–0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.08)[0]</td>
<td>(0.068)[0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>–0.15</td>
<td>–0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.098)[0.13]</td>
<td>(0.071)[0.028]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.46</td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.14)[0.001]</td>
<td>(0.14)[0.001]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.04</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.029)[0.16]</td>
<td>(0.027)[0.059]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuing Training Hours per Year, in Hundreds</td>
<td>0.03</td>
<td>–0.025</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.04)[0.45]</td>
<td>(0.029)[0.38]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.5</td>
<td>2.3</td>
<td>2.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.036)[0]</td>
<td>(0.049)[0]</td>
<td>(0.078)[0]</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>850</td>
<td>845</td>
<td>750</td>
<td>750</td>
</tr>
<tr>
<td>One Digit Industry Fixed Effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Notes: Pools all observations in 2003 and 2008 waves for establishments with at most 25 percent professional or managerial employees; 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.9: Regression of Average Log Wage on Individual HPWO Measures, Pooled Sample, Only Establishments with At Most 25 Percent Professional or Managerial Employees

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Log Wage (Std. Error)</th>
<th>[p-value]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Fraction in Self-Managed Teams</td>
<td>0.17</td>
<td>(0.098) [0.081]</td>
</tr>
<tr>
<td>Fraction in Job Rotation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Participating in FQM</td>
<td>0.05</td>
<td>(0.14) [0.73]</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.19</td>
<td>(0.083) [0.02]</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.16</td>
<td>(0.066) [0.015]</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>-0.55</td>
<td>(0.079) [0]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>-0.17</td>
<td>(0.1) [0.084]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.44</td>
<td>(0.14) [0.002]</td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.041</td>
<td>(0.029) [0.17]</td>
</tr>
<tr>
<td>Continuing Training Hours per Year, in</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hundreds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.39</td>
<td>(0.083) [0]</td>
</tr>
</tbody>
</table>

N: 764 768 759 768

Notes: Pools all observations in 2003 and 2008 waves for establishments with at most 25 percent professional or managerial employees; 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.10: Regression of Changes in Average Log Wage on Changes in Two Individual HPWO Measures, Balanced Panel, Restricted to Establishments with At Most 25 Percent Professional or Managerial Employees in 2003

<table>
<thead>
<tr>
<th>Variable (Differenced)</th>
<th>Average Log Wage (Std. Error)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Fraction in Self-Managed Teams</td>
<td>0.19 (0.098)[0.022]</td>
<td>0.18 (0.074)[0.013]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction in Job Rotation</td>
<td>−0.062 (0.071)[0.038]</td>
<td>−0.058 (0.069)[0.04]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.86 (0.36)[0.017]</td>
<td>0.67 (0.25)[0.008]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.39 (0.31)[0.02]</td>
<td>0.34 (0.25)[0.017]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>−0.11 (0.06)[0.08]</td>
<td>−0.064 (0.081)[0.43]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.04 (0.04)[0.32]</td>
<td>0.027 (0.046)[0.56]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuing Training Hours per Year, in Hundreds</td>
<td>−0.12 (0.054)[0.029]</td>
<td>−0.14 (0.055)[0.012]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>−0.3 (0.13)[0.022]</td>
<td>−0.27 (0.14)[0.047]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>−0.14 (0.11)[0.18]</td>
<td>−0.18 (0.1)[0.089]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.34 (0.2)[0.095]</td>
<td>0.41 (0.26)[0.11]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.034 (0.029)[0.24]</td>
<td>0.0086 (0.028)[0.76]</td>
<td>0.033 (0.028)[0.24]</td>
<td>0.0091 (0.03)[0.76]</td>
</tr>
</tbody>
</table>

N 174 123 177 124

Notes: Pools all establishments in the 2003 or 2008 waves that have at most 25 percent professional or managerial employees in 2003; 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.11: Regression of Log Hourly Entry Wage for Professionals and Managers on Summary HPWO Measure, Pooled Sample, Only Establishments Whose Most Common Occupation is Professional or Managerial

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Std. Err)</td>
<td>(p-value)</td>
<td>(Std. Err)</td>
<td>(p-value)</td>
</tr>
<tr>
<td>HPWO Summary Measure</td>
<td>0.2</td>
<td>0.15</td>
<td>0.075</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.004)</td>
<td>(0.068)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.16</td>
<td>0.17</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.004)</td>
<td>(0.089)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>-0.15</td>
<td>0.034</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.54)</td>
<td>(0.36)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>0.61</td>
<td>0.52</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.14)</td>
<td>(0.15)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>-0.72</td>
<td>-0.61</td>
<td>-0.61</td>
<td>-0.61</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.002)</td>
<td>(0.2)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.33</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.13)</td>
<td>(0.21)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.56</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.13)</td>
<td>(0.21)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>-0.006</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.01)</td>
<td>(0.052)</td>
<td>(0.099)</td>
</tr>
<tr>
<td>Continuing Training Hours per Year, in Hundreds</td>
<td>0.13</td>
<td>-0.024</td>
<td>-0.024</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.83)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.8</td>
<td>2.31</td>
<td>2.24</td>
<td>2.24</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.17)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>N</td>
<td>344</td>
<td>342</td>
<td>295</td>
<td>294</td>
</tr>
</tbody>
</table>

Notes: Pools all establishments in 2003 and 2008 waves for which the most common occupational category is professional or managerial workers; 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.12: Regression of Log Hourly Entry Wage for Professionals and Managers on Individual HPWO Measures, Pooled Sample, Only Establishments Whose Most Common Occupation is Professional or Managerial

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Hourly Entry Wage (Std. Error)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction in Self-Managed Teams</td>
<td>-0.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction in Job Rotation</td>
<td></td>
<td>0.024</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.23)</td>
<td>(0.92)</td>
<td></td>
</tr>
<tr>
<td>Fraction Participating in FQM</td>
<td></td>
<td></td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.09)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.21</td>
<td>0.2</td>
<td>0.12</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.098)</td>
<td>(0.041)</td>
<td>(0.098)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.079)</td>
<td>(0.12)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.03</td>
<td>0.051</td>
<td>-0.025</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.37)</td>
<td>(0.09)</td>
<td>(0.35)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.36)</td>
<td>(0.94)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health</td>
<td>0.47</td>
<td>0.45</td>
<td>0.56</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.12)</td>
<td>(0.12)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>-0.9</td>
<td>-0.83</td>
<td>-0.7</td>
<td>-0.77</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.25)</td>
<td>(0.25)</td>
<td>(0.23)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.32</td>
<td>0.34</td>
<td>0.35</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.096)</td>
<td>(0.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.13)</td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.61</td>
<td>0.6</td>
<td>0.58</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.15)</td>
<td>(0.15)</td>
<td>(0.15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>-0.013</td>
<td>-0.0079</td>
<td>0.0039</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.056)</td>
<td>(0.047)</td>
<td>(0.054)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.089)</td>
<td>(0.90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.054)</td>
<td></td>
</tr>
<tr>
<td>Continuing Training Hours per Year,</td>
<td>0.2</td>
<td>0.15</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>in Hundreds</td>
<td>(0.14)</td>
<td>(0.14)</td>
<td>(0.13)</td>
<td>(0.12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.28)</td>
<td>(0.32)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.25)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.38</td>
<td>2.36</td>
<td>2.21</td>
<td>2.27</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.15)</td>
<td>(0.14)</td>
<td>(0.15)</td>
</tr>
</tbody>
</table>

N 302 303 302 303

Notes: Pools all establishments in 2003 and 2008 waves for which the most common occupational category is professional or managerial; 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.13: Regression of Log High Wage for Professionals and Managers on Summary HPWO Measure, Pooled Sample, Only Establishments Whose Most Common Occupation is Professional or Managerial

<table>
<thead>
<tr>
<th>Variable</th>
<th>Log High Wage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) (Std. Error)</td>
</tr>
<tr>
<td>HPWO Summary Measure</td>
<td>0.23 (0.074) [0.002]</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.43 (0.15) [0.005]</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>-0.17 (0.41) [0.67]</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health</td>
<td>0.8 (0.15) [0]</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health</td>
<td>0.8 (0.15) [0]</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>-1 (0.25) [0]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.63 (0.35) [0.072]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.8 (0.17) [0]</td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>-0.015 (0.05) [0.76]</td>
</tr>
<tr>
<td>Continuing Training Hours per Year,</td>
<td>0.29 (0.096) [0.003]</td>
</tr>
<tr>
<td>in Hundreds</td>
<td>0.29 (0.096) [0.003]</td>
</tr>
<tr>
<td>Constant</td>
<td>3.26 (0.12) [0]</td>
</tr>
<tr>
<td>N</td>
<td>330</td>
</tr>
</tbody>
</table>

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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.14: Regression of Log High Wage for Professionals and Managers on Individual HPWO Measures, Pooled Sample, Only Establishments Whose Most Common Occupation is Professional or Managerial

<table>
<thead>
<tr>
<th>Variable</th>
<th>Log High Wage (Std. Error)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Fraction in Self-Managed Teams</td>
<td>−0.24</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Fraction in Job Rotation</td>
<td>−0.021</td>
<td>(0.24)</td>
</tr>
<tr>
<td>Fraction Participating in FQM</td>
<td>0.3</td>
<td>(0.094)</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.48</td>
<td>(0.23)</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.047</td>
<td>(0.54)</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>0.61</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>−1.19</td>
<td>(0.27)</td>
</tr>
<tr>
<td>Fraction Temporarily</td>
<td>0.61</td>
<td>(0.37)</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.82</td>
<td>(0.16)</td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>−0.019</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Continuing Training Hours per Year,</td>
<td>0.39</td>
<td>(0.12)</td>
</tr>
<tr>
<td>in Hundreds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.61</td>
<td>(0.14)</td>
</tr>
</tbody>
</table>

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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.15: Regression of Average Log Wage on HPWO Summary Measure, Establishment As A Whole, Pooled Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Log Wage</td>
<td>(Std. Error)</td>
<td>[p−value]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPWO Summary Measure</td>
<td>0.039</td>
<td>0.029</td>
<td>0.031</td>
<td>0.027</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.18</td>
<td>0.14</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>-0.082</td>
<td>0.064</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>0.33</td>
<td>0.15</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>-0.64</td>
<td>-0.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>-0.041</td>
<td>-0.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.56</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.22</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>-0.13</td>
<td>-0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.15</td>
<td>0.021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuing Training Hours</td>
<td>0.05</td>
<td>-0.0059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.63</td>
<td>2.39</td>
<td>2.41</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1404</td>
<td>1396</td>
<td>1202</td>
<td>1201</td>
</tr>
<tr>
<td>One Digit Industry Fixed Effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.16: Regression of Change in Average Log Wage on Change in HPWO Summary Measure, Establishment As A Whole, Balanced Panel Sample

<table>
<thead>
<tr>
<th>Variable (Differenced)</th>
<th>Change in Average Log Wage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>HPWO Summary Measure</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(0.043)[0.38]</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.044)[0.36]</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>−0.27</td>
</tr>
<tr>
<td></td>
<td>(0.34)[0.43]</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td>(0.082)[0.46]</td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.0037</td>
</tr>
<tr>
<td></td>
<td>(0.057)[0.95]</td>
</tr>
<tr>
<td>Continuing Training Hours per Year, in Hundreds</td>
<td>−0.15</td>
</tr>
<tr>
<td></td>
<td>(0.04)[0]</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>−0.29</td>
</tr>
<tr>
<td></td>
<td>(0.13)[0.028]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>−0.058</td>
</tr>
<tr>
<td></td>
<td>(0.1)[0.57]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>(0.26)[0.43]</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>−0.12</td>
</tr>
<tr>
<td></td>
<td>(0.11)[0.24]</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>−0.064</td>
</tr>
<tr>
<td></td>
<td>(0.075)[0.39]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(0.032)[0.23]</td>
</tr>
<tr>
<td></td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>(0.032)[0.26]</td>
</tr>
<tr>
<td></td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>(0.026)[0.32]</td>
</tr>
<tr>
<td></td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(0.026)[0.19]</td>
</tr>
<tr>
<td>N</td>
<td>253</td>
</tr>
<tr>
<td></td>
<td>252</td>
</tr>
<tr>
<td></td>
<td>188</td>
</tr>
<tr>
<td></td>
<td>177</td>
</tr>
</tbody>
</table>

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. Wages are in 2003 dollars. Weights used are 2003 establishment weights adjusted for nonresponse.
Table 2.17: Regression of Average Log Wage on Individual HPWO Measures, Establishment As A Whole, Pooled Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction in Self-Managed Teams</td>
<td>0.067</td>
<td>(0.071)</td>
<td>0.016</td>
<td>(0.043)</td>
</tr>
<tr>
<td>Fraction in Job Rotation</td>
<td>0.091</td>
<td>(0.082)</td>
<td>0.033</td>
<td>(0.043)</td>
</tr>
<tr>
<td>Fraction Participating in FQM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction in Problem Discussion Meetings</td>
<td>0.15</td>
<td>(0.058)</td>
<td>0.14</td>
<td>(0.056)</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.058</td>
<td>(0.13)</td>
<td>0.055</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>0.065</td>
<td>(0.062)</td>
<td>0.062</td>
<td>(0.067)</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>0.16</td>
<td>(0.055)</td>
<td>0.16</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>-0.63</td>
<td>(0.073)</td>
<td>-0.64</td>
<td>(0.074)</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>-0.045</td>
<td>(0.093)</td>
<td>-0.035</td>
<td>(0.094)</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.55</td>
<td>(0.091)</td>
<td>0.56</td>
<td>(0.091)</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.22</td>
<td>(0.1)</td>
<td>0.22</td>
<td>(0.1)</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>-0.13</td>
<td>(0.066)</td>
<td>-0.12</td>
<td>(0.066)</td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.016</td>
<td>(0.069)</td>
<td>0.016</td>
<td>(0.069)</td>
</tr>
<tr>
<td>Continuing Training Hours per Year, in Hundreds</td>
<td>0.054</td>
<td>(0.093)</td>
<td>0.059</td>
<td>(0.094)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.43</td>
<td>(0.085)</td>
<td>2.43</td>
<td>(0.081)</td>
</tr>
</tbody>
</table>

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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 2.18: Regression of Average Log Wage on HPWO Summary Measure, Pooled Sample, Establishment As A Whole, Restricted to Establishments with Some Bargaining Coverage

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPWO Summary Measure</td>
<td>0.15</td>
<td>0.13</td>
<td>0.081</td>
<td>0.047</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>0.16</td>
<td>0.1</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>0.088</td>
<td>-0.15</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>-0.19</td>
<td>-0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.083</td>
<td>-0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.91</td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>-0.4</td>
<td>-0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>0.18</td>
<td>-0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Training Hours, in Hundreds</td>
<td>0.015</td>
<td>-0.0099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuing Training Hours per Year,</td>
<td>0.17</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in Hundreds</td>
<td>(0.016)</td>
<td>(0.016)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.57</td>
<td>2.48</td>
<td>2.53</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>263</td>
<td>263</td>
<td>219</td>
<td>219</td>
</tr>
</tbody>
</table>

Notes: Pools all observations in 2003 and 2008 waves for which the fraction of workers covered under collective bargaining is greater than zero; 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Chapter 3

The Tenure-Size Relationship and the Age of Establishments

3.1 Introduction

It is well known that there is a positive association between the size of a firm (or establishment) and the average wage paid to its employees. (See, for example, Brown and Medoff (1989), for references to multiple sources and for reports of this correlation using multiple data sets.) Less frequently discussed, but still well recognized, is the positive correlation between the size of a firm and the average tenure of employees at the firm; we discuss more references below, but Brown, Hamilton and Medoff (1990) is one source that reports this correlation between tenure and firm size.

We demonstrate that the age of firms, or at least the age of establishments, plays a large part in this relationship. The key observation is that small firms have higher death rates than large firms and are thus, on average, younger than large firms. This has a mechanical effect on average tenure.\(^2\)

We note that we do not think that the age-size association necessarily explains the tenure-size relationship completely as, for example, quit rates are higher at smaller firms.

Many data sets do not contain information on all three of the variables tenure, firm size and age. A unique new data set, the California Business Survey (CBS), has all of this information, at the establishment level, and thus enables us to examine the interrelationships among these three quantities.

We briefly discuss a sampling of related work in the next section. In Section 3, we give a short description of the CBS data set we use. We present our results in Section 4 before summing up in the final section.

---

\(^1\)This chapter is joint work with Michael Reich.

\(^2\)We note that this line of reasoning has been observed before us by Brown, Hamilton and Medoff.
3.2 Theoretical Discussion and Related Literature

Numerous authors have reported on the tenure-size relationship. To give one example, Oi (1990) reports positive relationships between employee tenure and firm size for workers in U.S., Japan and Korea (with the exception of Korean professional workers). Using 1983 U.S. Current Population Survey data on private sector employees, he finds a mean tenure of 8.9 years for firms with 1000 or more workers and 4.2 years for firms with one to 24 workers.

As noted above, Brown, Hamilton and Medoff (1990) also report a positive tenure-size correlation. They actually discuss this relationship in the context of employee satisfaction, arguing that larger firms may have longer tenure because employees are happier at larger firms. But they note that the tenure-size association could also be due partly to larger firms being on average older, which they note is in turn caused by higher death rates among small firms and we would suggest is likely also caused by larger firms taking longer to grow.\(^3\) We posit that the size-age association, exactly in the way proposed by Brown, Hamilton and Medoff, actually can explain a large portion of the tenure-size gradient.\(^4\) Brown, Hamilton and Madoff report that quit rates are higher for smaller firms, and thus we do not necessarily conclude that all of the tenure-size relationship can be explained by firms’ size-age association. However, it is worth noting that a higher quit rate need not imply a lower average tenure as the resulting average tenure depends on the distribution of those who quit a firm.

Mumford and Smith (2004) explore the effects of individual characteristics and establishment characteristics on employee tenure, using linked employer-employee data from the United Kingdom. They find in general that the individual and employee characteristics each explain about the same proportion of the variation of tenure for individuals. They largely use establishment fixed effects, rather than characteristics, as controls. But in one case, they regress these calculated establishment fixed effects on workplace characteristics including establishment age, and find a statistically significant positive association (at the 5% level); in this regression, an establishment size variable is included but is not statistically significant. So these results are somewhat akin to ours and point in the same direction, but do not strike exactly at the same point. Curiously, workplace age is positive but not statistically significant when only “disproportionately female” workplaces are used, and negative and statistically significant when “disproportionately non-white” workplaces are used.

Davis, Haltianger and Schuh (1996) have data on employment levels at manufacturing plants in the U.S. They find a lower rate of gross job destruction at older plants, but they do not report any data related to job tenure.\(^5\) It thus seems likely

---

\(^3\)See page 56 of their work. See, for example, Evans (1987), Table II, for a source for firm survival decreasing with firm size.

\(^4\)We note that the mechanism is similarly suggested in a paper by Hope and Mackin (2007), but they similarly do not have the data to investigate it.

\(^5\)See pp.76-77 of their work.
that average job tenure is affected by establishment age in ways other than just by the mechanical effect that firms cannot, by definition, have employees with tenure any larger than their age.

However, one should bear in mind the fact that the data they use, the Longitudinal Research Database, does not record a job as being destroyed if it is replaced at the same plant in the following quarter that data is reported. The best one can do is create a composite job destruction total by summing up net job destruction numbers at each plant. While it seems quite plausible that the difference between this measure and true job destruction rates would not be influenced by establishment age, it is by no means clear.

Davis, Haltiwanger and Schuh also report that 23% of job destruction is accounted for by plant shutdowns. (This still refers to the manufacturing sector.) This gives some support to a theory that posits that much of the tenure-size relationship is due to higher death rates among small establishments.

Finally, we note that Brown and Medoff (2003) find a firm-level relationship between firm age and wages, but report that it disappears when one accounts for individual characteristics.

### 3.3 Data

The data set we use is a multiwave survey called the California Business Survey (CBS), which we briefly describe. The survey is a stratified sample of 1,080 establishments in California, over the period May 13 to October 22 of 2003, and a follow-up survey of 652 establishments in California, over the period April 29, 2008 to January 15, 2009. We refer throughout to the two waves as the 2003 and 2008 waves. It was conducted by the Survey Research Center at the University of California at Berkeley. Note that it is a survey of establishments, as opposed to firms. The CBS includes data on numerous characteristics of each sampled establishment, including, for example, number of employees, pay scale of the employees, occupational distribution of employees, employee tenure, hiring and training practices, general workplace practices, benefits, as well as detailed information on the most recently hired non-exempt and professional employees. 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 seems likely to be more accurate.

The sampling universe is the set of all nonprofit or business establishments in the state of California that have at least 5 employees, excluding public schools or universities, and agriculture, forestry and fishing enterprises, as well as any government
agencies. (Many of the details we describe here are taken from the CBS codebooks (Survey Research Center 2003 and 2009).) The sample was drawn from the database of information on California establishments maintained by Dun and Bradstreet.

The CBS 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 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 is 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 check for still being in business and not having moved out of California was made, reducing the numbers to 1,016. Then other eligibility criteria were checked. In addition to attempting to contact all of the 2003 survey participants, an additional sample of new firms was added to the 2008 survey in order to have a larger number of total observations. The sampling process was similar to the original 2003 sampling process described above. In 2008, two versions of the survey instrument were used: the original long version and a shortened version that omitted several questions from the longer instrument.

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+. Larger establishments were oversampled, with, for example, in 2003, 100% of the largest stratum being sampled, while in 2003 only 0.97% of the smallest establishments were sampled.

Finally, we briefly discuss our use of weights with the CBS. All weights that we use account for the rate of oversampling by stratum, and account for nonresponse. Some weights account 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. Generally, we use establishment weights in our regressions, unless mentioned otherwise.

In Table 3.1, we give some basic relevant descriptive statistics for the CBS. The estimates use data pooled from both waves of the survey, with a total of 1,732 ob-

---

8In 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.

9Note that classification into strata used D&B data, and in 2008, some establishments had no D&B data for establishment size or had a size of lower than 5 employees in the D&B data. As mentioned earlier, the surveyors checked with the establishments that all of these cases actually had at least 5 employees. Generally, in the paper we used the 2003 strata for these establishments.

10We 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.
servations. We highlight the means of some of the key variables. The average size of an establishment is around 44 employees. The tenure variable is reported by each establishment as an average over its full-time employees.\textsuperscript{11} The mean of this average, taken over all establishments, is 6.2 years. As with all of the variables in the table, the mean is calculated at the establishment-level using weights adjusted for nonresponse, and accounting for the stratification in the sample. So note that this mean is not weighted to adjust for the number of workers at each establishment.

Our measure of establishment age is the number of years that the establishment has been at a given site; an establishment may have moved from a previous location to the current site. It is not completely clear whether the most appropriate definition of establishment age would start at the date of a move or at the date of opening regardless of physical location. But even if one were to choose the latter, the quantity we use seems to be at least a reasonable proxy, as we suspect moves are not all that common. The average number of years on site, which we loosely refer to below as establishment age, is about 17 years.

The average hourly wage is about 18.5 dollars; this is in 2003 dollars, where the Consumer Price Index (CPI) has been used to adjust for inflation. 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.\textsuperscript{12} (Recall that the CBS does not include government establishments.)

### 3.4 Results

Before considering the effects of establishment age, we observe that average tenure is indeed positively associated with establishment size. We plot tenure against establishment size in Figure 3.1, using the pooled data. There is a general upward trend, although it is not extremely clear.

Having established that there is at least some correlation between tenure and establishment size in our data set, we turn to regression analysis to explore the issue. First, using the pooled sample, we consider three pairs of regressions. In each pair, we first regress average tenure on establishment size and potentially other controls; in the second regression of the pair, we add only the establishment age as a control and consider the effect on the estimate of the coefficient on establishment size.

Table 3.2 presents the results. Consider columns (1) and (2) first. In the first column, we regress average tenure on only establishment size and a constant. The estimated coefficient on establishment size is 2.12, while adding a control for the number of years the establishment has been on the site reduces this estimate to .97.

\textsuperscript{11}We note that the term full-time is not defined in the question, but an earlier question asking the number of part-time workers defines part-time as working less than 20 hours per week.

\textsuperscript{12}Our worker-weighted estimates are 12.3 percent in 2003 and 8.8 percent in 2008.
In the second column, the coefficient of establishment age is .092 and is statistically significant at a level of less than .05%, unsurprisingly giving strong evidence for a positive correlation of establishment age and average tenure. Columns (3) and (4) give a similar comparison, but now to both regressions we have included a control for the average log wage, which could be a significant factor in the relationship between tenure and establishment size. Again, the estimate of the coefficient on establishment size is reduced from 1.96 to 0.73; moreover, the latter estimate is significant only at the 15% level. The estimate of the coefficient for establishment age is nearly the same as in column (2) and is still strongly significant.13

Finally, in column (5), we regress average tenure on establishment size and numerous other variables as well as 1-digit industry fixed effects. These control variables include the fraction of workers in two occupational groups, the fraction of part-time employees, the fraction of temporary employees, the fraction having a college degree, whether or not the establishment offers a health insurance plan, and the fraction of workers at the establishment covered by collective bargaining. The estimate of the coefficient on establishment size is reduced, at 1.3 and is still statistically significant, although only at the 9% level. But when one adds the number of years the establishment has been on site, the point estimate of the coefficient on establishment size is in fact negative, at −.21. In columns (2), (4) and (6), the estimate of the coefficient on establishment age is remarkably consistent and always strongly statistically significant.

We can use the panel nature of the survey to investigate further. Note that there are only 358 observations in the balanced panel, and of course any regression within this subsample will have fewer observations due to unit nonresponse.

Table 3.3 presents the results. In the first column, we regress the change in average tenure from 2003 to 2008 on the change in establishment size, the change in average log wage, and a constant. 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.14 The inclusion of a constant allows for a constant shift in the log of replacement costs per employee across all establishments. The point estimate of the coefficient on establishment size is actually negative, but it is not statistically significant at any conventional level, as the standard errors are fairly large. In fact, the standard errors are large enough that a 95% confidence interval for the coefficient includes the estimate of 2.12 obtained in column (1) of Table 3.2. In column (2), we add the change in the number of years on the site to the regression. The estimate for establishment size changes little, but there is again, unsurprisingly, strong evidence for a positive association of average tenure and establishment age: the estimate is .069 and is statistically significant at the 0.4%

13The estimates in column (2) and (4) differ in the 5th digit.
14This approach would not capture all effects of unobserved characteristics, however, if the relationship between each establishment’s average tenure and the covariates changes across waves in a way other than by a simple constant additive shift in the tenure across all workplaces.
level. The value is very similar to the estimates on the variable in Table 3.2.

The third column excludes changes in establishment age, but includes numerous control variables. There are no one digit industry effects, nor controls for whether the workplace offers health insurance or for the fraction of employees covered by collective bargaining; these variables change values rarely across the years of the survey and thus estimates for these coefficients would have large standard errors. The fourth column adds the change in establishment age to this regression.

Again, the estimate for establishment size in column (3) is negative and not statistically significant. The magnitude is smaller than in columns (1) and (2), although not much should be concluded from this, as the standard errors are again large. The inclusion of establishment age changes the estimate for establishment size very little. There is still strong evidence for a positive correlation of establishment age and tenure, and at the same magnitudes seen before.

Regarding the control variables, the fraction of college workers is positively associated with tenure, and surprisingly the fraction of temporary workers is as well, although weakly so.

In summary, the balanced panel regression does not give any affirmation of the evidence found using the pooled sample for our proposal that the establishment size and tenure appear positively correlated only as a result of not controlling for establishment age. It could be the case that in this balanced panel regression, the establishment fixed effects control sufficiently for any confounding variables, so that establishment age has no additional effect on estimates of the coefficient on establishment age. Further, one might speculate that a larger sample could rule out any substantial positive correlation.

3.5 Conclusion

Above we have presented some evidence that the observed relationship between the size of an establishment and the average tenure there can perhaps be largely explained by the age of the establishment. Smaller firms may have higher death rates than larger firms, and also it may take firms time to grow; both of these mechanisms imply an association between firm size and firm age. This, together with the mechanical relationship between tenure and firm age, may explain the apparent tenure-size correlation. We find fairly strong evidence from the pooled sample to support this explanation, but results from the balanced panel are inconclusive.

Recall that there is evidence that suggests that firm size is related to tenure for reasons other than just firm age. Brown, Hamilton and Madoff (1990) report that quit rates are higher for smaller firms, which indicates that firm age does not account for all of the observed tenure-size association. Nevertheless the evidence we have presented strongly suggests that firm age accounts for a very substantial portion of the relationship.
Figure 3.1: Average Tenure by Establishment Size Category

Notes: Establishment size category is indicated in the first line of text below each bar. Standard errors and number of observations used for each calculation are also indicated below each bar, denoted by "SE" and "N", respectively. Weights adjust for nonresponse, and the standard error calculations account for the stratified nature of the sample. Establishments in both waves of the panel that changed strata across waves are assigned to the stratum in the first wave. Estimates are clustered at the level of the establishment (across waves). Average years of tenure is the average numbers of years worked at the establishment among full-time employees.
<table>
<thead>
<tr>
<th>Quantity</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Std. Error)[Std. Dev.]</td>
</tr>
<tr>
<td>Establishment Size</td>
<td>43.83</td>
</tr>
<tr>
<td>Average Years of Tenure</td>
<td>6.186</td>
</tr>
<tr>
<td>Years at Site</td>
<td>16.949</td>
</tr>
<tr>
<td>Fraction Covered by Bargaining Agreement</td>
<td>0.031</td>
</tr>
<tr>
<td>Offer Health Insurance Plan</td>
<td>0.824</td>
</tr>
<tr>
<td>Average Hourly Wage</td>
<td>18.546</td>
</tr>
<tr>
<td>Average Log Hourly Wage</td>
<td>2.67</td>
</tr>
<tr>
<td>Hourly Entry Wage</td>
<td>13.191</td>
</tr>
<tr>
<td>High Hourly Wage</td>
<td>22.318</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>0.178</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.053</td>
</tr>
<tr>
<td>Fraction College</td>
<td>0.306</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.282</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>0.365</td>
</tr>
<tr>
<td>Number of Sites Firm Operates</td>
<td>90.121</td>
</tr>
</tbody>
</table>

Notes: Uses pooled sample. Mean values of variables in the table are taken at the establishment level, where the same establishment is treated as two observations if it appears in both waves. Estimates are clustered at the level of the establishment (across waves). Standard errors are in parentheses while standard deviations are in brackets. Weights adjust for nonresponse, and the calculations account for the stratified nature of the sample. Average years of tenure is the average numbers of years worked at the establishment among full-time employees. Years at site is the number of years the firm has “been in operation at [the] site.” It is answered in buckets or specified as a number by the respondent if more than 10 years; in the case of buckets, the midpoint is used, while in the case of given numbers, the reports were hand-coded. Wage figures are in 2003 dollars. Average log hourly wage is calculated using midpoints of ranges provided for percentages of each establishment’s workers having hourly wages in a given range; the Pareto distribution is used to calculate a point estimate for the highest range. Hourly high and entry wage are for the establishment’s largest employee category, among the following choices: (i) professional and managerial employees, (ii) clerical employees, (iii) sales employees and (iv) manual labor and blue collar employees.
Table 3.2: Regression of Average Tenure on Establishment Size and Length at Site, Pooled Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years at Site</td>
<td>0.092</td>
<td>0.092</td>
<td>0.088</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>2.12</td>
<td>0.97</td>
<td>1.96</td>
<td>0.73</td>
<td>1.3</td>
<td>−0.21</td>
</tr>
<tr>
<td></td>
<td>(0.69)</td>
<td>(0.002)</td>
<td>(0.53)</td>
<td>(0.068)</td>
<td>(0.004)</td>
<td>(0.51)</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>0.47</td>
<td>0.63</td>
<td>0.61</td>
<td>0.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.5)</td>
<td>(0.34)</td>
<td>(0.48)</td>
<td>(0.19)</td>
<td>(0.66)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td>(1.07)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.99)</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>(0.78)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.16)</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>2.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td>(0.95)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.033)</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>(0.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.48)</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>−0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−0.049</td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.81)</td>
</tr>
<tr>
<td>Offer Health Insurance Plan</td>
<td>1.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>(0.56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.029)</td>
</tr>
<tr>
<td>Fraction Covered by Collective Bargaining</td>
<td>3.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.73</td>
</tr>
<tr>
<td></td>
<td>(1.42)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.03)</td>
</tr>
<tr>
<td>Constant</td>
<td>6.09</td>
<td>4.58</td>
<td>4.84</td>
<td>2.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.26)</td>
<td>(0.001)</td>
<td>(0.037)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N 1688 1642 1625 1585 1379 1347

Notes: Standard errors are in parentheses and p-values are in brackets. A p-value of 0 denotes that the p-value is less than 0.005. Pools all observations in the 2003 and 2008 wave; observations are at the establishment-year level. Estimates are clustered at the level of the establishment (i.e., across waves). Weights adjust for nonresponse, and the standard error calculations account for the stratified nature of the sample. Establishments in both waves of the panel with a missing stratum in the second wave are assigned to the stratum in the first wave. Average years of tenure is the average numbers of years worked at the establishment among full-time employees. Average log of hourly wage is in 2003 dollars. 1-digit fixed effects do not include a fixed effect for the Standard Industrial Classifications (SIC) starting with 9, as these are in the category of public administration and thus not in our sample, with the exception of one erroneously classified observation.
Table 3.3: Regression of Changes in Average Tenure on Changes in Establishment Size and Age and Other Variables, Balanced Panel

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years at Site</td>
<td>0.069</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.024)[0.004]</td>
<td>(0.022)[0.002]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>−1.07</td>
<td>−0.97</td>
<td>−0.19</td>
<td>−0.2</td>
</tr>
<tr>
<td></td>
<td>(1.8)[0.55]</td>
<td>(1.53)[0.53]</td>
<td>(2.31)[0.93]</td>
<td>(1.87)[0.91]</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>0.41</td>
<td>−0.092</td>
<td>0.45</td>
<td>−0.15</td>
</tr>
<tr>
<td></td>
<td>(0.91)[0.65]</td>
<td>(0.8)[0.91]</td>
<td>(0.9)[0.62]</td>
<td>(0.91)[0.87]</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.42</td>
<td>1.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.11)[0.84]</td>
<td>(2.11)[0.45]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>−1.63</td>
<td>−1.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.58)[0.3]</td>
<td>(1.72)[0.48]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>1.44</td>
<td>1.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.93)[0.46]</td>
<td>(1.93)[0.47]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>2.49</td>
<td>2.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.57)[0.11]</td>
<td>(1.5)[0.14]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>4.02</td>
<td>3.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.21)[0.069]</td>
<td>(2.06)[0.08]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.91</td>
<td>0.79</td>
<td>0.94</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>(0.38)[0.017]</td>
<td>(0.37)[0.032]</td>
<td>(0.39)[0.017]</td>
<td>(0.37)[0.042]</td>
</tr>
<tr>
<td>N</td>
<td>311</td>
<td>294</td>
<td>280</td>
<td>265</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses and p-values are in brackets. A p-value of 0 denotes that the p-value is less than .0005. Uses all observations in the balanced panel, that is all establishments in both the 2003 and 2008 waves. Weights adjust for nonresponse, and the standard error calculations account for the stratified nature of the sample. Weights and strata are assigned based on their value in the first wave. Average years of tenure is the average numbers of years worked at the establishment among full-time employees.
Chapter 4
Determinants of Voluntary Quit Rates

4.1 Introduction

Quit rates are an area of key interest in Labor Economics, arising, for example, in matching models, and of interest in their own right.\(^1\) In this article, we look at whether certain employer policies affect voluntary quit rates. By voluntary quit rates, we mean quits in which the worker chooses to either leave the labor force or in which the worker moves to another job.\(^2\) There is little work on these policies and their relationship with quit rates, as the data required is fairly rare. We study the relationship with a new survey of California establishments.

Fairris (Fairris 2004) has provided a notable contribution to this literature. He focuses on the effects of internal labor market practices on quit rates; some examples of such practices include: (i) priority being given to current workers over outside applicants when hiring for an open position above entry level, and (ii) the importance of seniority when making a promotion from within to an open position. Key results he gives include finding that (i) is associated with increased worker quits, while (ii) is associated with decreased worker quits. He also investigates job ladders; we discuss these results more below.

We look at similar questions. Additionally, the data set we analyze has more measures of worker turnover rates at the establishment level than have been generally studied in the past, allowing us to assess more carefully the reasons workers leave. At this point, we feel it is important to emphasize (as does Fairris, in a fairly similar way) that although there is good reason to believe that there is causation going in the direction from our various labor market practices to worker quit rates, our results are certainly only associations. Naturally it could be the case that establishments

---

\(^1\)See, for example, (Burdett and Mortensen 1998).

\(^2\)Voluntary quit rates are our focus, but we do investigate other turnover rates to some extent.
might first make note of high quit rates, and subsequently take up these internal labor market policies in order to reduce worker quits. Nevertheless, as it would be very difficult, if not impossible, to randomly assign labor market practices to firms, evidence of associations is clearly quite valuable.

There are of course numerous characteristics of establishments that could affect worker quit rates. In the next section, we discuss some of these, and consider what theory suggests about their relationship with quit rates. In Section 4.3, we describe the survey data we use. In Section 4.4, we detail the key results, while the following section compares to other work. Section 4.6 concludes.

4.2 Potential Determinants of Quit Rates: Theory

There are many possible factors associated with voluntary quit rates at establishments. Throughout this section, we focus on voluntary quits, but we examine other turnover rates in the empirical section.\(^3\) We first consider health insurance coverage, and whether or not it is provided by an establishment.

In the United States at least, one might expect health insurance coverage to be associated with lower quit rates. Certainly if one is considering the effect of health insurance coverage given a fixed wage, one would expect coverage to be associated with reduced quit rates, as the health insurance would provide higher utility to a worker. However, if one does not control for wage, it could be that health insurance is provided as a compensating differential associated with a lower wage, as in the theory of Rosen (Rosen 1993). In this case, it is not as clear what one would expect the effect of health insurance coverage on quit rates to be. Indeed, for turnover generally, this question has been extensively studied, and with varying results. See Gruber and Madrian (Gruber and Madrian 2004) for a theoretical framework and survey of results, and Madrian (Madrian 2006) for a more recent survey of empirical work. So, if one does not control for wages, the prediction is unclear about the association between health insurance coverage and quit rates.

We now consider another policy: whether or not workers at an establishment are in a union, or what fraction of them are. The more relevant quantity is the fraction of workers that are covered by a collective bargaining agreement, as this measures directly the fraction of workers whose wages and benefits might be raised by the presence of a union. In this setting, the competitive labor market model is likely not valid, depending on the union’s bargaining power, and workers may be able to appropriate rents due to the bargaining agreement. One might expect that workers could also secure rents other than wages—e.g., unions could bargain for improved working conditions or job security. Thus certainly one would expect lower quit rates to be associated with higher fractions of worker bargaining coverage, even after controlling for wages.

\(^3\)Multiple strands in this section follow Fairris (Fairris 2004) to some extent.
We turn to a few internal labor market policies. Consider first how much an establishment prioritizes hiring workers from within when filling open positions above entry level, as opposed to hiring from an outside applicant pool. Given a fixed wage, it seems possible that this policy could lead to lower quit rates, if workers value an environment in which they are given a leg up in being hired for an open position at the establishment. Alternatively, as Fairris (Fairris 2004) suggests, it could be the case that an unpleasant atmosphere of competitiveness might be fostered by such a policy, if it is well-known that open positions are more likely to be filled by present employees. If this is the dominant effect, the policy could lead to higher quit rates. Additionally, as Fairris also posits, such a policy might lead firms to promote higher-skilled workers than they might otherwise hire from outside, since these workers could be more accurately screened than outside workers; this would in turn result in a workforce with higher average skills, which would consequently have more outside job opportunities and thus higher quit rates.

There is another mechanism by which prioritizing internal hiring could lead to higher voluntary quits. Suppose that prioritizing internal hiring were to take the particular form of a so-called “up-or-out” policy, in which firms select certain workers to place on long and stable job ladders, while assigning the remaining workers to either remain indefinitely at particular job levels or leave the firm. Brown et al. (2010), for example, found evidence consistent with the presence of such policies in the semiconductor, software and financial services industries, and similar evidence in some cases in the trucking and retail foods industries. If a firm operates with an up-or-out policy and the number of workers selected for long job ladders is much larger than the number of those not selected, then we would expect this policy to result in increased voluntary quit rates, as those not selected would leave for other firms at which they might have a chance to advance their careers.

In summary, it is not clear what the predicted relationship between internal promotion priority and quit rates should be, as there are possible mechanisms that could cause negative or positive associations.

As in the discussion above, in general there is some question of whether internal labor market policies influence quits only via their effect on the human capital of the workforce, or whether there is another channel. To this end, we control for human capital variables in our regressions, although we do not have exceptionally strong controls of this type.

Consider now a question asking establishments how much they value seniority when they make promotion decisions, given that the open position is to be filled from within. In this case, theory suggests that the jobs of workers with longer tenures have larger nonwage value to them, in that their chance of being promoted increases with tenure. Thus the expected future wage of workers increases with tenure, and therefore so does the present value of the current job to a worker.\footnote{This assumes of course that workers do not discount the future entirely.} So we expect that
quit rates would be negatively associated with the weight the establishment puts on seniority when promoting internally.

Now we consider three measures concerning job ladders. In the data we investigate, the job ladder length measured in time refers to the average number of years it takes for a worker to rise from the lowest to highest position of the establishment’s most common occupational category (assuming there are at least two levels). These categories are: (i) professional or managerial; (ii) sales; (iii) clerical; and (iv) manual labor or blue collar.\(^5\) We also have another way of measuring job ladder length. We define the job ladder length measured in job levels as the number of job levels in the most common occupational category. Finally, we define the wage growth from bottom to top of the job ladder as the difference between the high pay and entry pay in the most common employee category, taken as a fraction of the entry pay.

Fixing the first two job ladder measures, job level length and time length, and also fixing the wage, we would expect the last job ladder measure, wage growth, to have a negative effect on quit rates, as the expected future wage is of course larger if the growth is larger. On the other hand, we would expect the number of job levels in a job ladder to have a negative relationship with worker quit rates: given a fixed time to climb the job ladder and wage growth from bottom to top, we would expect workers to prefer a shorter number of levels in the job ladder, as this implies a larger rate of annual growth of wages. However, as Fairris points out, it could be the case that workers might feel an increased sense of improvement by climbing a longer ladder, and may prefer knowing they would not need to incur larger search costs in the future in order to find another job. Finally, given fixed wage growth and a fixed number of job levels in the ladder, we would expect the time to climb the job ladder to be positively associated with quit rates, although with a potential negative association for the same reasons as for the number of job levels.

Andersson et al. (2008) provide an interesting study that investigates job ladders and highlights the role that heterogeneity might play in the effects of internal labor market policies. They examine human resource management practices for firms in the electronics industry, emphasizing how these practices may affect worker productivity differently, depending on whether the firm is in an industry with relatively fast or slower technical change. They use longitudinal linked employer-employee data to investigate these questions, and find, for example, that productivity at firms with high investment in R&D (Research and Development) is larger at firms with lower turnover, performance incentives, and multiple ports of entry (which allows firms to hire workers with needed skills rather than train them). They conclude generally that high R&D firms do better using performance-based human resource systems, while low R&D firms benefit from a more bureaucratic system in which for example they do not choose workers for a specific track at an early stage, but in which the

---

\(^5\)In survey questions, the second and third categories are separate in the first wave, but combined in the second.
system nevertheless has the flexibility to enable them to hire from outside at multiple position levels.

4.3 Data

4.3.1 Overview

The data set we use is a panel survey called the California Business Survey or CBS (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 CBS 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.6

The CBS 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.7

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 CBS 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

---

6The 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.

7The CBS 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.
were selected at a higher rate. This sample (2806 in 2003) was checked to ensure it met the eligibility criteria, namely 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. 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. The survey designers therefore wrote an abbreviated version that omitted a number of questions from the original instrument.

Table 4.1 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+. 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 4.1 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).

Finally, we briefly discuss our use of weights with the CBS. All weights that we use account for the rate of oversampling by stratum, and account for nonresponse. 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,

---

8In 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.

9Classification 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.


11As 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.
4.3.2 Key Variables and Descriptive Statistics

In Table 4.2, we give detailed descriptions for those key variables whose brief name in the regression tables might not completely convey their meanings. In particular we look at various turnover rates, which will be our central response variables. Note that these are all annual rates. The central quit rate we study is the number of workers who leave an establishment for another job or because they leave the labor force, as a fraction of the total number of employees at the establishment. We also present some labor market policies, for example health insurance eligibility and the importance of seniority when promoting an employee within the establishment.

Table 4.3 gives summary statistics for some relevant variables. We present the mean, standard error (in parentheses) and standard deviation (in brackets) for each variable, for the pooled sample of both waves. Recall that the survey is stratified, with different sampling rates for each stratum. Each calculation accounts for these establishment weights, as well as the stratified nature of the sample. Means here are not weighted by the number of workers in each establishment; thus here means are to be interpreted as taken over establishments, and not as means for all workers.

There is an average establishment-level turnover rate of 32.1 percent per year, with a 18.1 percent voluntary quit rate, whether for another job, or for education, retirement or family reasons. The turnover rate due to being fired for performance, in 2003 only, is 14.1 percent (not reported). It is worth considering the context of the employment situation at each time. The simple mean of the (seasonally adjusted) U.S. unemployment rate from the Current Population Survey, during the six months in 2003 during which the survey took place, was 6.13 percent; the analogous number for the 2008 survey is 6.4 percent. The 2003 number was generally higher than during the surrounding years; for example the annual unemployment rates in 2002 and 2004 were 5.8 percent and 5.5 percent, respectively. Officially, the preceding recession ended in November 2001, although many have referred to the recovery after the

---

12Note that percentages reported in Table 4.1 are simple unweighted percentages.

13In many cases the descriptions use the original wording directly from the original codebooks (Survey Research Center 2003b, Survey Research Center 2009). Some of the variables are available only in the 2003 survey. Also, as mentioned above, the classification of occupational categories varies between survey waves; in 2008, the clerical and sales occupations were grouped together.

14In 2003, within total turnover, there were also questions asked about transfer, layoffs and job eliminations. The 2003 survey asked how much attrition there was in the previous year, as well as the number of workers who left in the previous year for various reasons. In participants’ responses, the sum of the workers who left for various reasons does not necessarily add up to the response for the total number of workers who left, perhaps due to reporting error or due to workers leaving for an alternative reason not mentioned in the survey. In fact, the sum over all individual turnover rates in 2003 is 41.0 percent, yet the 2003 total turnover rate is 35.8 percent.

15See http://www.nber.org/cycles.html
2001 recession as a jobless recovery. In 2008, although the unemployment number was not much larger than in 2003, the survey took place as a recession that officially started in December 2007 grew gradually worse, with a 7.7 percent unemployment rate by January 2009.

The mean size of establishments in the survey is 43.8. The mean of the establishment average hourly wage is around 18.5 dollars per hour.\(^\text{16}\) The establishments’ largest occupational category, on average, is blue collar and manual labor workers, with 36.5 percent of the establishment’s workforce, followed by professional and managerial workers, with the clerical and sales categories trailing.

The mean establishment has only 3.1 percent of workers covered under 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.\(^\text{17}\) (Recall that the CBS does not include government establishments.)

Respondents at the mean establishment state that they consider giving preference to current employees when filling a vacancy above entry level as somewhere between very important (a value of 3) and extremely important (a value of 4). They view using seniority when making decisions about which employees to promote as between moderately important (a value of 2) and very important (a value of 3). Establishments do have job ladders: in the most common occupational category, whether professional, sales, clerical or blue collar, the mean number of job levels is 3.13. The mean time to climb this job ladder, for those establishments for which the ladder contains at least two job levels, is 4.7 years.

### 4.4 Results

Table 4.4 presents regressions of the quit rate on the health insurance eligibility rate for full time workers. 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.\(^\text{18}\) In the first column, we regress on only the health insurance

\[^{16}\text{The latter is calculated as a mean over all employees in the establishment, using responses to questions about what fraction of employees' wages are in particular wage categories. This mean, calculated for each establishment, is then averaged over all establishments to give the final reported number.}\]

\[^{17}\text{Our worker-weighted estimates are 12.3 percent in 2003 and 8.8 percent in 2008. This places our estimates in line with national numbers: according to the Bureau of Labor Statistics, 9.1 percent of U.S. private sector wage and salary workers in non-agricultural industries had contracts covered by a union or an employee association in 2003, and 8.5 percent in 2008. See BLS Series LUU0204917300, retrieved from http://data.bls.gov/cgi-bin/srgate .}\]

\[^{18}\text{All regressions account for the establishment weights, which are adjusted for nonresponse, and for the stratified nature of the sample.}\]
eligibility rate, and we add the average log wage in the second column. In this manner, we can test the theory about the effect of the health insurance eligibility rate with and without controlling for the wage. The third and fourth columns add controls for establishment size, the fraction of employees covered by collective bargaining, the fractions of part-time and temporary workers, the fraction of workers with a college degree and the fraction of workers in two occupational categories. Again, the average log wage is excluded in the third column, but included in the fourth. The last two columns are similar to the third and fourth columns except that now we include eight indicators as controls for one digit Standard Industrial Classification (SIC) codes, and remove the constant.

Health insurance eligibility is negatively associated with leaving the establishment, whether we control for the establishment’s average log wage or not. In all cases, the association is statistically significant at the .4 percent level or lower. Eligibility is associated with reducing the number of employees leaving the workplace voluntarily by about 11 percent per year. Thus an establishment’s changing its policy from offering health insurance to no full time employees to a policy offering it to all full time employees is associated with a change of more than half the mean (establishment-level) quit rate of 18.1 percent per year. (See Table 4.3.) It is worth bearing in mind that the mean level of full time health insurance eligibility is 75.8 percent, so in practice this is not applicable to most employers, even if the association is in fact causal.

In the regressions in Table 4.4, the wage is negatively associated with quit rates, as expected (although not always in a strongly statistically significant fashion). So is the establishment size, as has been noted in other settings.\footnote{See, for example, p. 57 of (Brown, Hamilton and Medoff 1990).} We discuss bargaining coverage shortly, but note that surprisingly, the fraction of workers part-time or temporary do not show associations at the statistical significance one might expect. The point estimates for the associations with the fraction of college workers are both positive and negative, whereas one would expect a negative result. On the other hand, quit rates are negatively associated with the fraction of professional or managerial employees, and statistically significant at the 8 percent level or lower. Quit rates are also negatively associated with the fraction of blue collar or manual labor workers. The reference group here is sales and clerical workers; perhaps these rates are influenced largely by comparison with clerical workers.

Table 4.4 also allows us to examine the fraction of employees covered by collective bargaining, in columns (3)-(6). The point estimates are negative, but not statistically significant at any conventional level. A bivariate regression of the quit rate on this fraction and a regression including also the average log wage yield slightly larger point estimates (unreported), but similarly not statistically significant estimates. However, examination of graphs highlights one particular outlier, with a quit rate of more than 300 percent a year. When removed, regressions (both bivariate and with multiple
controls as in column (6) of Table 4.4) yield negative estimates of about \(-0.08\) (or more negative) that are statistically significant at the level of \(0.0015\) or lower.\textsuperscript{20} Thus there is suggestive evidence that the quit rate is negatively associated with increased bargaining fraction.\textsuperscript{21}

Table 4.5 focuses on hiring and promotion preferences, and job ladders. The first three columns present regressions of the quit rate on these practices, while column 4 adds the various controls that were used in Table 4.4. The final column adds one digit industry fixed effects (and removes the constant term).

Giving preference to current employees over outside workers when filling a vacancy is associated positively with quit rates. However, this association is not robustly statistically significant; in fact, it is very weak in the bivariate regression, as can be seen in the first column. In the last two columns however, the estimate is significant at traditional levels, with a \(p\)-value of \(0.033\) or lower.

Treating seniority as important in promotion, when hiring from within, does not appear to be associated with quit rates; the point estimates in the bivariate regression in column 2 are positive while the point estimates in the final two columns are negative, but not at any conventional level of statistical significance.

As far as job ladders, the length of the ladder (the number of job levels in the most common occupational category) and the time to climb the ladder show no clear association with quit rates. The change in pay between the bottom and top of the wage ladder, however, is negatively and significantly associated with quit rates, although the level of statistical significance is as high as 8 percent. To get a sense of the size of the estimates in the final two columns, increasing this change in pay by one standard deviation (67.3 percent) is associated with lowering the annual quit rate by an absolute amount of 1.7 percentage points: that is, a reduction of 9.7 percent of the mean quit rate.

To investigate these relationships further, we can make use of the panel nature of the data, by estimating a regression of differences in the quit rate, 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 control to some extent for unobserved characteristics of each establishment. However, these regressions use very few observations, as few as 82 once one includes multiple controls. Thus we

\textsuperscript{20}As usual, one should be cautious removing observations based on the value of the outcome variable.

\textsuperscript{21}One might be concerned that removing this outlier would change some of the other results in the paper. The outlier is from the 2008 wave, so to check whether this is the case, we need only perform the regressions that use both waves of data but with this outlier removed. These are the regressions reported in Tables 4.4 and 4.5. As we have already discussed, removing this one outlier changes the estimates on the coefficient of bargaining coverage in regressions analogous to those reported in Table 4.4; however, the only other estimate that is substantially different is for the coefficient of the fraction of temporary workers, which although now much smaller, is not statistically significant with or without the outlier in the sample. For the regressions in Table 4.5, the results are not substantially changed.
do not report the results in a table. We nevertheless note that none of the above-mentioned results, e.g. for the hiring policy or the wage growth variable, obtain for these balanced panel regressions. These facts should be interpreted in light of the size of the standard errors: the standard errors in the balanced panel regression associated with, for example, the estimates of the coefficient of the hiring preference variable, are in each specification at least as large as the respective point estimates in Table 4.5. This bolsters our inclination to conclude that we do not have enough power to make any conclusions from the balanced panel.

We turn to finer measures of turnover, which are available only in the 2003 survey. In Table 4.6, we consider the portion of the quit rate stemming from only those quits in which workers leave the establishment for another job.22 The first three columns focus on the hiring and promotion policies, and job ladder characteristics, while the fourth column includes multiple controls for establishment characteristics, and the last column includes one digit industry fixed effects.

Giving preference to current employees is positively associated with increased quits of this type, though the correlation is much smaller (and not statistically significant) in the bivariate regression. In unreported regressions, we find no evidence on the other hand that the turnover rate due to leaving the labor force is associated with this hiring preference. If prioritizing internal hiring leads to a more competitive atmosphere, then this atmosphere might affect quits to leave the labor force and quits to take another job equally. So this competitive atmosphere channel does not appear consistent with these results.

However, consider the above-mentioned story in which prioritizing internal hiring leads to more screening of workers; in turn, this leads to more skilled workers in the establishments, who thus have greater outside opportunities, and thus quit more often to take other jobs. These greater skills would not seem as likely to affect the quit rate to leave the workforce and could even negatively affect this quit rate, if greater skills led to increased wages and thus led to higher utility through working relative to the utility of leisure. Thus the screening channel is consistent with our results.

Also recall the up-or-out policy discussed above, in which only some workers are selected for stable career tracks within an establishment. In this setting, many of those not selected may quit to look elsewhere for positions offering career advancement. One would expect these workers to quit to go to other jobs; we would not expect them to leave the workforce due to the up-or-out policy. So if the workers not selected form the large majority of the establishment, then one would expect to see an increase in the quit rate to go to other jobs, but no effect on the quit rate to leave the workforce. Thus an explanation of the effect of internal hiring priority on quit rates via the channel of an up-or-out policy is, in addition to the screening story, consistent with our results.

---

22Recall that the quit rate considered above is based on quits either to go to another job or to leave the labor force.
Estimates of the coefficients on job ladder length in years and the fractional change in wage from bottom to top of the ladder are negative and statistically significant in column (3). However, the job ladder length estimates are only very weakly significant in the final two columns. The fractional change in wage is also weakly negatively associated with quits to leave the labor force, but there is no similar association for either of the job ladder length variables (unreported). So, although the results are weak, the change in pay as one climbs the job ladder appears to be negatively associated with either type of voluntary quit. The job ladder length in years, on the other hand, is negatively associated with only quits to accept another job (but very weakly so), which is in line with the intuition that only those who plan to stay in the workforce would prioritize having a longer job ladder.

Other estimates of note in Table 4.6 include a weakly significant (at the 8 percent level) negative association of quits to accept another job with the fraction of full-time workers eligible for health insurance, and (weak) negative associations of these quits with the log wage and establishment size. The association of the rate of such quits with bargaining coverage is negative and statistically significant at the level of 5 percent or lower, perhaps due to low motivation to change jobs if one has a unionized job and its potential nonwage benefits, e.g., added job security. The estimates for the coefficient in a regression of the quit rate to leave the labor force on bargaining coverage are negative, but weaker and much smaller in magnitude. So having a unionized job may affect one’s decision to leave the labor force, but not as much as it affects one’s choice of job. Finally, the fraction of blue collar workers is negatively associated with quits to accept other jobs, perhaps because fewer jobs are available to these workers.

In Table 4.7, we report regressions of a turnover rate for an involuntary type of turnover, namely the annual rate at which workers leave the establishment because their performance was unsatisfactory. Again, this uses data only from the 2003 survey. The specifications are as in Table 4.6: the first three columns focus on the hiring and promotion policies, and job ladder characteristics, while the fourth column includes multiple controls for establishment characteristics and the last adds one-digit industry fixed effects.

Both (i) giving preference to current employees in hiring and (ii) prioritizing seniority in promotion decisions are negatively associated with this turnover rate, although only very weakly in the bivariate regressions in the first two columns. Perhaps such policies encourage more worker effort, in the spirit of efficiency wage theories, and thus less firing. Alternatively, the policies might simply be more common at establishments where workers are more skilled.

Time to climb the job ladder is negatively associated with turnover due to unsatisfactory performance, but only once one controls for multiple establishment characteristics; it may be the case that in establishments with long job ladders, there is more incentive for employees to work hard to prove themselves. On the other hand, we see no such negative association for the number of job levels in the job ladder.
Eligibility for health insurance is weakly negatively associated with this type of turnover, as perhaps retaining health insurance is a motivator for better performance. The fraction of temporary workers is negatively associated with unsatisfactory performance turnover rates, perhaps unexpectedly. Finally, and surprisingly, the average log wage shows no relationship with firing due to unsatisfactory performance.

Table 4.8 reports regressions of the quit rate due to accepting another job, on the monthly employee copayment for the establishment’s health insurance plan.\textsuperscript{23} The first column reports results from a regression of this type of quit rate on the copayment and the average log wage, for all establishments. The second column reports the same regression, but now restricted to the sample of establishments with 100 percent health insurance eligibility for full time employees. It is these establishments at which one might expect a larger effect of the monthly copayment. In the third column, we continue to restrict to this sample, and add multiple establishment controls. We keep the same sample and add one digit industry fixed effects in the final column.

Given a fixed wage, the size of the employee copayment is positively associated with the quit rate to go to another job, at the 10 percent level or lower (and at the 5 percent level or lower once multiple controls are included). Given a fixed wage, a higher copayment is in effect reduced compensation for labor, and thus one would expect higher quits as a result, perhaps especially to seek a job with a better total compensation package. As is also expected, Table 4.8 shows that the log wage is negatively associated with these quits; the association is statistically significant with a \( p \)-value of 4.7 percent or lower.

We now mention some additional specifications we investigated but do not report in the paper. In many of the regressions we have reported, the establishment size is negatively related to the various quit rates. Moreover, in Chapter 3, we showed that much of the relationship between establishment size and tenure can be explained by the association between establishment size and establishment age. Given the relationship between tenure and quit rates, one may be concerned that establishment size might simply be proxying for establishment age in many of the reported regressions. As a check, we performed all of the reported regressions, but including establishment age as an additional regressor wherever we had included establishment size. All results are substantially the same.

Training hours could be endogeneous with respect to the quit rate; for example, firms might reduce their training hours in response to a high quit rate. Given such concerns, we have not included measures of training in our analysis. However, one may be interested to see if including training as a regressor might have some effect on estimates of the partial correlations of the internal labor market policies. Thus, we performed the regressions in Tables 4.5, 4.6 and 4.7, but with measures of both new training hours and continuing training hours as additional controls (for those specifications that already include many controls). The results are basically the same,

\textsuperscript{23}Again, this uses data only from the 2003 survey.
4.5 Comparison with Other Work

It is of interest to compare our results with the earlier work of Fairris (Fairris 2004), as it is similar in many ways, although different in many others. First, there are differences with the data sets. Fairris uses a data set gathered by Osterman (Osterman 1994), which is a nationwide survey of establishments, called the National Survey of Establishments (NSE). It differs crucially from the CBS sample we use, in that it only surveys establishments with at least 50 employees, whereas the lower limit for the CBS is 5 workers. Also, the NSE is limited to for-profit establishments, whereas the CBS can include nonprofits but does exclude the agriculture, forestry and fishing sectors and public schools and government, as mentioned above. Of course the CBS is also limited to California.

Another key difference is that many of the questions Fairris uses pertain to the “core employee” group, that is, the “largest group of non-supervisory, non-managerial workers at this location who are directly involved in making the product or providing the service” at the establishment. (See p. 6 of (Massachusetts Institute of Technology and The Center for Survey Research 1992).) Most of the key quantities Fairris investigates refer to the core employee group, including quit rates, the level of preference given to current employees in filling a vacancy, prioritization of seniority in hiring, and questions about job ladders. Quit rates, preference for hiring current employees and seniority prioritization in the CBS refer to all of the establishments’ employees. On the other hand, questions about job ladders in the CBS refer to the most common occupational category and thus may be somewhat similar to the quantities Fairris uses, except for the stipulation in the NSE that those quantities refer to non-managerial and non-supervisory workers.

The phrasing of the quit rate question in the sample Fairris uses is “What was your voluntary quit rate among [the core employee group] for the last 12 months?” (See p. 6 of (Massachusetts Institute of Technology and The Center for Survey Research 1992).) Note that the survey used in that paper does not have the finer breakdown that the CBS has, so all of Fairris’s results relate to this quit rate.

Fairris finds that giving preference to current employees when filling a vacancy is associated with increased quit rates, while prioritizing seniority when promoting from within is associated with decreased quit rates. We find similar results for hiring preferences to current employees, although our results are not robust. We do not find any clear evidence of associations with prioritizing seniority in promotion. Concerning

24 The analogous estimates are significant at the 9 and 14 percent levels for internal hiring priority and at the 11 and 9 percent levels for wage growth.
job ladders, Fairris only uses wage growth and the time to climb the job ladder; the data set he uses does not have information about the number of job levels in the ladder. He finds negative associations with quit rates for wage growth (as do we, albeit not robustly so), but also negative associations of quit rates with time to climb the job ladder, which we do not find.

4.6 Conclusion

We have looked at the relationship of quit rates with several labor market policies and with properties of establishments. We have found some evidence for negative associations of voluntary quit rates with policies such as offering health insurance and bargaining coverage.\textsuperscript{25}

We focus more on less frequently studied policies, including the prioritization of hiring workers from within over outside workers, and the prioritization of seniority when filling a job from within. We do not identify any evidence for either a positive or negative association of quit rates with prioritizing seniority in promotion. However, we find from pooled regressions that prioritizing hiring workers from within is positively associated with increased quit rates. Regressions using a balanced panel with fixed effects for each establishment do not confirm this result; however, the sample sizes for this balanced panel are very small and hence standard errors are sufficiently large that the lack of corroboration could be due solely to a lack of statistical power.

We also find a weak positive association of prioritizing hiring workers from within with the quit rate to take another job, but not with quits to leave the labor force. This combination of results is consistent with two potential explanations. In one story, prioritizing internal hiring leads to more screening of workers; in turn, this leads to more skilled workers in the establishments, who thus have greater outside opportunities, and thus quit more often to take other jobs. This combination of results could also occur if establishments had an “up-or-out” policy, in which employers select relatively few workers for long job ladders; any workers not given promotions or increased wages would quit for other firms that might offer them the potential for career advancement.

We also investigate job ladders and their relationship with quit rates. We find no association with quit rates for either the job ladder length in job levels or the job ladder length in years. However, a measure of the relative change in pay from entry to the highest level of the most common occupational category has a weak negative relationship with quit rates.

The turnover rate due to unsatisfactory performance is negatively associated with both prioritizing hiring from within and prioritizing seniority in promotion. As we have noted, this could be due to employees exerting more effort in response to these policies, akin to an efficiency wage story.

\textsuperscript{25}The evidence is somewhat weak in the latter case.
Finally, we find positive associations of monthly employee health insurance copayments with quit rates due to accepting another job, given a fixed wage, as one would expect, given that increased copayments are essentially decreases in compensation.

We emphasize again that the research design does not allow us to interpret these associations as implying causality. However, it is extremely unlikely that one could find natural experiments, and thus such associations are important as the only tool available to study these complex and interesting relationships.

Throughout, some of the results mentioned were not robust; we note that this could be an issue of statistical power, but it could be a genuine failure of the result to hold. Further investigation is certainly needed.
Table 4.1: Sampling Results and Response Rates in the CBS

<table>
<thead>
<tr>
<th></th>
<th>2003</th>
<th>2008 Panel Sample</th>
<th>2008 New Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampled Establishments</td>
<td>2806</td>
<td>1080</td>
<td>1072</td>
</tr>
<tr>
<td>2008 Panel Sample Still in Business in California</td>
<td></td>
<td>1016</td>
<td></td>
</tr>
<tr>
<td>Number Meeting Eligibility Criteria</td>
<td>2200</td>
<td>868</td>
<td>849</td>
</tr>
<tr>
<td>Interview Completed</td>
<td>1080</td>
<td>358</td>
<td>294</td>
</tr>
<tr>
<td>Response Rate (Unweighted)</td>
<td>49.1%</td>
<td>41.2%</td>
<td>34.6%</td>
</tr>
</tbody>
</table>
Table 4.2: Definitions of Key Variables

<table>
<thead>
<tr>
<th>Turnover Rates</th>
<th>General Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Turnover Rate</strong></td>
<td><strong>Fraction Temporary</strong></td>
</tr>
<tr>
<td>Ratio of number of employees leaving establishment in the last year to total</td>
<td>Fraction of employees that are temporary workers, hired for a specific task,</td>
</tr>
<tr>
<td>number of employees at the establishment</td>
<td>project or limited time frame</td>
</tr>
<tr>
<td><strong>Quit Rate</strong></td>
<td><strong>Fraction Part-Time</strong></td>
</tr>
<tr>
<td>Fraction of employees leaving site in last year due to accepting another</td>
<td>Fraction of employees that work less than 20 hours per week</td>
</tr>
<tr>
<td>job or voluntarily leaving labor force for education, family reasons or</td>
<td></td>
</tr>
<tr>
<td>retirement</td>
<td></td>
</tr>
<tr>
<td><strong>Quit Rate due to Accepting Another Job</strong></td>
<td><strong>Fraction College Degree</strong></td>
</tr>
<tr>
<td>Fraction of employees leaving site in last year due to accepting another job</td>
<td>Fraction of employees that have a college degree</td>
</tr>
<tr>
<td>(2003 Only)</td>
<td></td>
</tr>
<tr>
<td><strong>Turnover Rate due to Leaving Labor Force</strong></td>
<td><strong>Fraction Professional or Managerial</strong></td>
</tr>
<tr>
<td>Fraction of employees leaving site in last year because they voluntarily left</td>
<td>Fraction of employees that are in professional or managerial occupations</td>
</tr>
<tr>
<td>labor force for education, family reasons or retirement (2003 Only)</td>
<td></td>
</tr>
<tr>
<td><strong>Turnover Rate due to Unsatisfactory Performance</strong></td>
<td><strong>Fraction Blue Collar</strong></td>
</tr>
<tr>
<td>Fraction of employees leaving site in last year because their performance was</td>
<td>Fraction of employees that are in manual labor or blue collar occupations</td>
</tr>
<tr>
<td>unsatisfactory (2003 Only)</td>
<td></td>
</tr>
<tr>
<td><strong>General Variables</strong></td>
<td><strong>Fraction Clerical</strong></td>
</tr>
<tr>
<td><strong>Fraction Temporary</strong></td>
<td>Fraction of employees that are in clerical occupations (2003 Only)</td>
</tr>
<tr>
<td><strong>Fraction Part-Time</strong></td>
<td><strong>Fraction Sales</strong></td>
</tr>
<tr>
<td><strong>Fraction College Degree</strong></td>
<td>Fraction of employees that are in sales occupations (2003 Only)</td>
</tr>
<tr>
<td><strong>Fraction Professional or Managerial</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Fraction Blue Collar</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Fraction Clerical</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Fraction Sales</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Labor Market Policies</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Fraction under Collective Bargaining</strong></td>
<td>Fraction of employees covered by a collective bargaining agreement</td>
</tr>
<tr>
<td><strong>Fraction Full Time Eligible for Health Insurance</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Continuing Training Hours per Year</strong></td>
<td></td>
</tr>
<tr>
<td><strong>New Training Hours</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Hiring Preference to Current Employees</strong></td>
<td></td>
</tr>
<tr>
<td>How important is it to give preference to current employees when filling a</td>
<td></td>
</tr>
<tr>
<td>vacancy above entry level; scale ranging from 4 (extremely important) to 1</td>
<td></td>
</tr>
<tr>
<td>(slightly important)</td>
<td></td>
</tr>
<tr>
<td><strong>Importance of Seniority in Promotion</strong></td>
<td></td>
</tr>
<tr>
<td>How important is it to use seniority as a criterion when filling a vacancy</td>
<td></td>
</tr>
<tr>
<td>with a current employee; scale ranging from 4 (extremely important) to 1 (</td>
<td></td>
</tr>
<tr>
<td>slightly important)</td>
<td></td>
</tr>
<tr>
<td><strong>Job Ladder Length (Job Levels)</strong></td>
<td>Number of job levels in the establishment’s most common employee category</td>
</tr>
<tr>
<td><strong>Job Ladder Length (Years)</strong></td>
<td></td>
</tr>
<tr>
<td>Years to rise from bottom to top of the establishment’s most common employee</td>
<td></td>
</tr>
<tr>
<td>category, for those establishments for which the ladder contains at least two</td>
<td></td>
</tr>
<tr>
<td>job levels</td>
<td></td>
</tr>
<tr>
<td><strong>Bottom to Top Wage Growth</strong></td>
<td>Change in hourly pay between entry level and highest level in the establishment’s</td>
</tr>
<tr>
<td></td>
<td>most common employee category, as fraction of entry pay</td>
</tr>
</tbody>
</table>
Table 4.3: Descriptive Statistics for Key Variables (Pooled Sample)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>(Std. Error)</th>
<th>[Std. Dev.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Turnover Rate</td>
<td>0.321</td>
<td>(0.03)</td>
<td>[0.6]</td>
</tr>
<tr>
<td>Quit Rate</td>
<td>0.181</td>
<td>(0.01)</td>
<td>[0.25]</td>
</tr>
<tr>
<td>Establishment Size</td>
<td>43.83</td>
<td>(2.81)</td>
<td>[15.73]</td>
</tr>
<tr>
<td>Average Hourly Wage</td>
<td>18.546</td>
<td>(0.52)</td>
<td>[10.26]</td>
</tr>
<tr>
<td>Average Log Hourly Wage</td>
<td>2.67</td>
<td>(0.02)</td>
<td>[0.43]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.306</td>
<td>(0.01)</td>
<td>[0.27]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.053</td>
<td>(0.01)</td>
<td>[0.27]</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>0.178</td>
<td>(0.01)</td>
<td>[0.24]</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.282</td>
<td>(0.01)</td>
<td>[0.25]</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>0.365</td>
<td>(0.02)</td>
<td>[0.33]</td>
</tr>
<tr>
<td>Fraction Covered by Collective Bargaining</td>
<td>0.051</td>
<td>(0.01)</td>
<td>[0.15]</td>
</tr>
<tr>
<td>Fraction Full Time Eligible for Health Insurance</td>
<td>0.758</td>
<td>(0.02)</td>
<td>[0.4]</td>
</tr>
<tr>
<td>Hiring Preference to Current Employees</td>
<td>3.203</td>
<td>(0.04)</td>
<td>[0.89]</td>
</tr>
<tr>
<td>Importance of Seniority in Promotion</td>
<td>2.181</td>
<td>(0.05)</td>
<td>[1.01]</td>
</tr>
<tr>
<td>Job Ladder Length (Job Levels)</td>
<td>3.133</td>
<td>(0.06)</td>
<td>[2.9]</td>
</tr>
<tr>
<td>Job Ladder Length (Years)</td>
<td>4.7</td>
<td>(0.21)</td>
<td>[3.9]</td>
</tr>
<tr>
<td>Bottom to Top Wage Growth</td>
<td>0.673</td>
<td>(0.03)</td>
<td>[0.78]</td>
</tr>
<tr>
<td>New Training Hours</td>
<td>65.591</td>
<td>(4.11)</td>
<td>[76.85]</td>
</tr>
<tr>
<td>Continuing Training Hours per Year</td>
<td>34.648</td>
<td>(3)</td>
<td>[55.5]</td>
</tr>
<tr>
<td>Establishment Age</td>
<td>16.949</td>
<td>(0.79)</td>
<td>[17.18]</td>
</tr>
</tbody>
</table>

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. Standard deviations are in brackets. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 4.4: Regression of Quit Rate on Health Insurance Eligibility for Full Time Employees, Pooled Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction FT Eligible for Health Insurance</td>
<td>-0.11  (0.037)</td>
<td>-0.11  (0.038)</td>
<td>-0.12  (0.035)</td>
<td>-0.11  (0.036)</td>
<td>-0.12  (0.035)</td>
<td>-0.11  (0.036)</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>-0.076  (0.026)</td>
<td>-0.066  (0.030)</td>
<td>-0.057  (0.041)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>-0.079  (0.026)</td>
<td>-0.073  (0.025)</td>
<td>-0.068  (0.025)</td>
<td>-0.06  (0.023)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>-0.036  (0.051)</td>
<td>-0.025  (0.051)</td>
<td>-0.03  (0.055)</td>
<td>-0.025  (0.053)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>0.12    (0.053)</td>
<td>0.078    (0.062)</td>
<td>0.098    (0.064)</td>
<td>0.062   (0.066)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.13    (0.11)</td>
<td>0.12     (0.11)</td>
<td>0.12     (0.11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>-0.02   (0.048)</td>
<td>0.015    (0.048)</td>
<td>-0.014   (0.046)</td>
<td>0.021   (0.046)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>-0.16  (0.059)</td>
<td>-0.13    (0.061)</td>
<td>-0.13    (0.057)</td>
<td>-0.11   (0.061)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>-0.16   (0.052)</td>
<td>-0.18    (0.053)</td>
<td>-0.13    (0.053)</td>
<td>-0.15   (0.055)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.29    (0.045)</td>
<td>0.47     (0.075)</td>
<td>0.36     (0.058)</td>
<td>0.52    (0.11)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N 1285 1243 1232 1197 1231 1196
One Digit Industry No No No No Yes 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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 4.5: Regression of Quit Rate on Various Internal Labor Market Policies, Pooled Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hiring Preference to Current Employees</td>
<td>0.016</td>
<td>0.033</td>
<td>0.032</td>
<td>(0.013)[0.19]</td>
<td>(0.014)[0.022]</td>
</tr>
<tr>
<td>Importance of Seniority in Promotion</td>
<td>0.0016</td>
<td>−0.0075</td>
<td>−0.0088</td>
<td>(0.013)[0.9]</td>
<td>(0.013)[0.48]</td>
</tr>
<tr>
<td>Job Ladder Length (Job Levels)</td>
<td>−0.0027</td>
<td>0.0013</td>
<td>0.0016</td>
<td>(0.0022)[0.22]</td>
<td>(0.0022)[0.54]</td>
</tr>
<tr>
<td>Job Ladder Length (Years)</td>
<td>−0.0084</td>
<td>−0.0005</td>
<td>0.0003</td>
<td>(0.0039)[0.031]</td>
<td>(0.0034)[0.09]</td>
</tr>
<tr>
<td>Bottom to Top Wage Growth, in Hundreds</td>
<td>−0.036</td>
<td>−0.027</td>
<td>−0.026</td>
<td>(0.016)[0.024]</td>
<td>(0.014)[0.056]</td>
</tr>
<tr>
<td>Fraction FT Eligible for Health Insurance</td>
<td>−0.15</td>
<td>−0.15</td>
<td>(0.057)[0.004]</td>
<td>(0.053)[0.006]</td>
<td>(0.053)[0.006]</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>−0.11</td>
<td>−0.091</td>
<td></td>
<td>(0.043)[0.014]</td>
<td>(0.049)[0.067]</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>−0.064</td>
<td>−0.056</td>
<td></td>
<td>(0.025)[0.011]</td>
<td>(0.026)[0.03]</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>−0.079</td>
<td>−0.084</td>
<td></td>
<td>(0.039)[0.042]</td>
<td>(0.04)[0.034]</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>0.071</td>
<td>0.051</td>
<td></td>
<td>(0.073)[0.48]</td>
<td>(0.073)[0.48]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>−0.013</td>
<td>−0.017</td>
<td></td>
<td>(0.013)[0.32]</td>
<td>(0.013)[0.19]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.0033</td>
<td>−0.0081</td>
<td></td>
<td>(0.056)[0.95]</td>
<td>(0.054)[0.88]</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>−0.13</td>
<td>−0.14</td>
<td></td>
<td>(0.082)[0.13]</td>
<td>(0.091)[0.14]</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>−0.2</td>
<td>−0.17</td>
<td></td>
<td>(0.071)[0.006]</td>
<td>(0.075)[0.028]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.13</td>
<td>0.18</td>
<td>0.26</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.037)[0.001]</td>
<td>(0.029)[0]</td>
<td>(0.032)[0]</td>
<td>(0.15)[0]</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1419</td>
<td>1417</td>
<td>908</td>
<td>798</td>
<td>797</td>
</tr>
<tr>
<td>One Digit Industry Fixed Effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

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. Weights used are establishment weights adjusted for nonresponse. Weights are combination of 2003 and 2008 weights.
Table 4.6: Regression of Quit Rate due to Accepting Another Job on Various Internal Labor Market Policies, 2003 Wave Only

<table>
<thead>
<tr>
<th>Variable</th>
<th>(Std. Error)</th>
<th>(p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Hiring Preference to Current Employees</td>
<td>0.013</td>
<td>0.031</td>
</tr>
<tr>
<td>Importance of Seniority in Promotion</td>
<td>0.0013</td>
<td></td>
</tr>
<tr>
<td>Job Ladder Length (Job Levels)</td>
<td>0.0002</td>
<td>0.0014</td>
</tr>
<tr>
<td>Job Ladder Length (Years)</td>
<td>−0.0099</td>
<td>−0.0027</td>
</tr>
<tr>
<td>Bottom to Top Wage Growth</td>
<td>−0.041</td>
<td>−0.035</td>
</tr>
<tr>
<td>Fraction FT Eligible for Health Insurance</td>
<td>−0.12</td>
<td>−0.11</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>−0.094</td>
<td>−0.084</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>−0.039</td>
<td>−0.031</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>−0.093</td>
<td>−0.096</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>−0.11</td>
<td>−0.11</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>−0.062</td>
<td>−0.058</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>−0.037</td>
<td>−0.061</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>−0.16</td>
<td>−0.17</td>
</tr>
<tr>
<td>Fraction Sales</td>
<td>−0.011</td>
<td>−0.053</td>
</tr>
<tr>
<td>Constant</td>
<td>0.092</td>
<td>0.13</td>
</tr>
<tr>
<td>N</td>
<td>913</td>
<td>909</td>
</tr>
<tr>
<td>One Digit Industry Fixed Effects</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Notes: Regressions use 2003 data only. Standard errors are in parentheses. p-values are in brackets. A p-value of 0 indicates that p < 0.0005 holds. Weights used are establishment weights adjusted for nonresponse.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Turnover Rate due to Unsatisfactory Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Std. Error) [p-value]</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>(5)</td>
</tr>
<tr>
<td>Hiring Preference to Current Employees</td>
<td>$-0.032$</td>
</tr>
<tr>
<td></td>
<td>(0.02) [0.11]</td>
</tr>
<tr>
<td>Importance of Seniority in Promotion</td>
<td>$-0.015$</td>
</tr>
<tr>
<td></td>
<td>(0.016) [0.35]</td>
</tr>
<tr>
<td>Job Ladder Length (Job Levels)</td>
<td>$-0.0008$</td>
</tr>
<tr>
<td></td>
<td>(0.0027) [0.78]</td>
</tr>
<tr>
<td>Job Ladder Length (Years)</td>
<td>$-0.0043$</td>
</tr>
<tr>
<td></td>
<td>(0.0057) [0.45]</td>
</tr>
<tr>
<td>Bottom to Top Wage Growth</td>
<td>$-0.093$</td>
</tr>
<tr>
<td></td>
<td>(0.097) [0.34]</td>
</tr>
<tr>
<td>Fraction FT Eligible for Health Insurance</td>
<td>$-0.097$</td>
</tr>
<tr>
<td></td>
<td>(0.056) [0.085]</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(0.06) [0.53]</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>$-0.03$</td>
</tr>
<tr>
<td></td>
<td>(0.02) [0.15]</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>$-0.0041$</td>
</tr>
<tr>
<td></td>
<td>(0.002) [0.9]</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>$-0.047$</td>
</tr>
<tr>
<td></td>
<td>(0.06) [0.44]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>$-0.13$</td>
</tr>
<tr>
<td></td>
<td>(0.059) [0.025]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>$-0.015$</td>
</tr>
<tr>
<td></td>
<td>(0.072) [0.83]</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>(0.084) [0.099]</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>(0.078) [0.013]</td>
</tr>
<tr>
<td>Fraction Sales</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>(0.073) [0.15]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(0.099) [0.014]</td>
</tr>
<tr>
<td></td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(0.071) [0.02]</td>
</tr>
<tr>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(0.1) [0.021]</td>
</tr>
<tr>
<td></td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(0.18) [0.3]</td>
</tr>
<tr>
<td>N</td>
<td>927</td>
</tr>
<tr>
<td></td>
<td>922</td>
</tr>
<tr>
<td></td>
<td>624</td>
</tr>
<tr>
<td></td>
<td>602</td>
</tr>
<tr>
<td></td>
<td>602</td>
</tr>
</tbody>
</table>

Notes: Regressions use 2003 data only. Standard errors are in parentheses. p-values are in brackets. A p-value of 0 indicates that $p < 0.0005$ holds. Weights used are establishment weights adjusted for nonresponse.
Table 4.8: Regression of Quit Rate due to Accepting Another Job on Monthly Employee Copayment, 2003 Wave Only

<table>
<thead>
<tr>
<th>Variable</th>
<th>Quit Rate due to Accepting Another Job</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Std. Err.) [p-value]</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Monthly Employee Copayment, in Hundreds of Dollars</td>
<td>0.022 (0.013)[0.096]</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>−0.072 (0.035)[0.042]</td>
</tr>
<tr>
<td>Establishment Size, in Thousands</td>
<td>−0.051 (0.018)[0.004]</td>
</tr>
<tr>
<td>Bargaining Coverage</td>
<td>−0.053 (0.041)[0.19]</td>
</tr>
<tr>
<td>Fraction Part Time</td>
<td>0.01 (0.001)[0.87]</td>
</tr>
<tr>
<td>Fraction Temporary</td>
<td>0.17 (0.11)[0.12]</td>
</tr>
<tr>
<td>Fraction College Degree</td>
<td>0.033 (0.047)[0.48]</td>
</tr>
<tr>
<td>Fraction Professional or Managerial</td>
<td>−0.12 (0.071)[0.099]</td>
</tr>
<tr>
<td>Fraction Blue Collar</td>
<td>−0.12 (0.066)[0.07]</td>
</tr>
<tr>
<td>Fraction Sales</td>
<td>−0.19 (0.069)[0.006]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.3 (0.1)[0.003]</td>
</tr>
</tbody>
</table>

N: 729 598 585 585
Restricted to Establishments with 100 percent Health Insurance Eligibility: No Yes Yes Yes
One Digit Industry Fixed Effects: No No No Yes

Notes: Regressions use 2003 data only. Note that columns (2)-(4) restrict to establishments with 100 percent health insurance eligibility for full time workers. Standard errors are in parentheses. p-values are in brackets. A p-value of 0 indicates that p < 0.0005 holds. Weights used are establishment weights adjusted for nonresponse.
Chapter 5

Can Econometric Methods Compensate for the Absence of a Randomized Control Group? The RAND Health Insurance Experiment

5.1 Introduction

This paper considers whether standard econometric techniques might lead to similar estimates to those obtained from randomized experiments, following the groundbreaking work of Lalonde (1986). Lalonde investigates how econometric evaluations of the effect of a training program on incomes would compare with the estimate obtained from a true, randomized social experiment designed to evaluate this effect. He analyzes the National Supported Work Demonstration (NSW), a federally sponsored training program in the 1970s, in which applicants were randomly assigned to a treatment group (the training program) or control group. One can easily calculate the difference in means of income in the year after training for the treatment and control groups to estimate the average effect of the training program. Lalonde compares this result to estimates an econometrician might obtain if there had been a group in the training program but there was no control group—i.e., if a researcher was forced to use another similar group as a substitute for the true control group. Lalonde finds that many estimates one would obtain, using multiple econometric techniques, either vary widely from the “true” training effect (that is, the mean income difference in the social experiment) or are estimates that an econometrician would be unlikely to choose as a reliable estimate.

These results are certainly very important, as they cast some doubt on a large
portion of econometric work. In his work, Lalonde claims that the only way to test econometric methods is by comparing the estimates one obtains with results from genuine experimental data. He calls for other researchers to look at other social experiments. We turn to one very famous experiment, mentioned by Lalonde in fact, and consider how estimates from econometric methods compare to results of this experiment. At the same time, we delve more deeply into the setup and outcomes of the experiment.

This social experiment is the so-called RAND Health Insurance Experiment (HIE). From 1974 to 1982, the RAND Corporation, supported by funding from the federal Office of Economic Opportunity and later by the Department of Health, Education and Welfare, ran a social experiment in which 8,254 individuals were assigned randomly to various health plans. One prominent way in which the health plans varied was in terms of their coinsurance rate, which is the percentage that an insured person has to pay for their medical treatments (at least up to a certain maximum expenditure limit). For example, some were assigned to a plan in which all medical care was free (a 0% coinsurance rate), while others were assigned to a plan in which, up to a certain limit, they were required to pay 95% of their medical expenses (a 95% coinsurance rate).

The central purpose of the HIE was to estimate the effect of health insurance plans on the demand for medical treatment, an important concern in the context of contemporaneous consideration of a national universal health program in the U.S. The HIE also estimated the effects on health of participants in the various health plans, in addition to many other subexperiments.

Our basic methodology is similar to Lalonde’s: we try to arrive at estimates a researcher would obtain if free insurance were offered to one set of participants, but no group of participants was randomly assigned to receive insurance with a higher coinsurance rate. In other words, we might think of this situation as if we were in a different universe in which there is a treatment program but no randomized control group; in this case, we might refer to this as a natural experiment. To take the place of an analogous control group, we turn to data from a national health survey conducted around the time of the HIE. The key idea is that within this comparison group many of the individuals are either insured at a higher coinsurance rate than 0% or completely uninsured. Indeed, according to a report on health insurance in 1975, “most insured persons had to pay directly for a substantial part of their total health care costs because of the exclusions, restrictions, and limitations that characterize[d] private health insurance policies.” Moreover, roughly 15% of workers were uninsured

---

1The Department of Health, Education and Welfare later became the Department of Health and Human Services.

2However, there was some criticism of the experiment’s design toward that end. See Hester and Leveson (1974) and Veney (1974).

3See p. 3, Mueller (1977). In the context of private insurance, 80% of hospital costs were estimated to be covered by insurers, with the rest mostly out-of-pocket, with less than 50% of costs
in 1979. This is far from an ideal control group, but of course this is an important part of Lalonde’s argument—often there is no ideal control group available. The national health survey we use is the Second National Health and Nutrition Examination Survey (NHANES II), conducted between 1976 and 1980. (See McDowell et al. (1981) for details.) We consider a few estimators, using the NHANES II data as a substitute for a control group, to test these econometric methods.

We analyze two health outcomes for which comparable data is available in the NHANES II survey. We choose one of these health outcomes because it is one of two key results RAND researchers found demonstrating a beneficial effect of free insurance for the average participant. We also look at another very similar health outcome. Experimental results indicate no effect on this latter health outcome, but we include it as a weak test of the econometric techniques, noting that, at a significance level of 5%, a technique with a decision rule that randomly assigns a result to be significant 5% of the time should of course find no effect 95% of the time.

The RAND experiment had varying refusal rates for the different plans. Moreover, the refusal rates for different plans seem to have varied by the values of certain characteristics of the individuals, such as the size of their family.

Thus we examine pre-experiment and demographic characteristics of participants in the control and treatment groups, as a test of the validity of the randomization. We also inspect more closely the two key results RAND researchers found. Finally, as family size appears to play a role in differential plan refusal, and because of the obvious interest in income as a health and economic factor, we also investigate some interesting outcome variables by family size and household income categories.

The plan of the paper is as follows. In the next section we discuss some of the related literature. In section 5.3, we describe both data sets, with particular emphasis on the RAND HIE data set. We then turn to describing our methods, in section 5.4.1 those for comparing the nonexperimental econometric estimates with the experimental estimates, and in section 5.4.2 our methods for examining the validity of the randomization and of some of the RAND researchers’ results. Section 5.4.3 describes how we investigate the outcomes by family size and income categories. In Section 5.5, we report the results found using each of the above methods. The final section concludes.

covered for “physicians’ services”, which includes “surgery, in-hospital physicians’ visits, and X-ray and laboratory examinations.” (p. 5 and p. 3, respectively, of Mueller (1977).)


5The sampling process took place in two stages. In the first, all potential participants were interviewed about their past use of medical and dental care, and demographic characteristics. Then the experiment designers determined eligibility, and randomly assigned eligible families to plans; at this point, the families could choose whether or not to enroll.

6The ideal analysis in such a case would be an intention-to-treat analysis, but the RAND researchers were not of course able to measure health outcomes of nonparticipants at the end of the trial.
5.2 Related Literature

We briefly note some of the related literature on these topics. The work of Lalonde (1986) touched off a series of papers by several authors discussing econometric methodology. First, the contributions of Leamer (1983, 1985) pointing out problems with, and suggesting improvements in, the econometric use of regressions should be noted as important work preceding Lalonde’s. After Lalonde’s work, several groups of researchers applied alternative nonexperimental estimation methods to the data set that he used (without the experimental controls), in an effort to produce estimates close to the experimental estimates. Heckman and Hotz (1989) offered a specification test for choosing among several nonexperimental estimators; applying their tests to certain subcategories of the sample, they found estimates close to the nonexperimental estimates. Dehejia and Wahba (1999) applied the propensity score of Rosenbaum and Rubin (1983) to the same data set and arrived at estimates close to the experimental estimates.

Many studies arose from the RAND HIE; we mention only a few. A key publication is a 1993 compendium of the results, by Joseph Newhouse and many of the RAND researchers. The central article showing that higher coinsurance rates do lead to more demand for medical services is Manning et al. (1987). Brook et al. (1983) report improvements from the free plan on vision and on blood pressure for low-income participants. Keeler et al. (1985) report a beneficial effect of the free plan on diastolic blood pressure, and produce evidence on the mechanism for the effect. Newhouse et al. (1987) respond to a critique of the findings of increased health expenditure on the free plan, and in so doing, discuss the topic of differential refusal rates for the plans.

5.3 Data

We use data from the RAND Health Insurance Experiment (Newhouse (1974–1982)) as well as data from the Second National Health and Nutrition Examination Survey (U.S. Department of Health and Human Services (1976–1980)), which we refer to as NHANES II. We describe each data set, with emphasis on the RAND data set.

5.3.1 RAND HIE

The RAND data set consists of some information for 26,148 people, with 24,340 of these being given an extensive interview and questionnaire before enrollment (which we loosely refer to together as the baseline interview), which asked about demographic data and self-reported health status and health attitudes. For one site (Dayton, Ohio), the questionnaire took place in two steps. See p. 8 of Opskalski et al. (1987).
health insurance plan, using random assignment to determine the offered plan. In all, 8,254 people were assigned to one of several health insurance plans.\footnote{Note that this includes some participants who were enrolled without initial interviews, largely adoptees or infants born to families on a health insurance plan. See pp. 1-2, Polich et al. (1986).} Every non-self-supporting and otherwise eligible member of a given family was enrolled in the same plan as all other family members.\footnote{See p. 13, Newhouse (1993).}

There were 15 different insurance plans offered, including one plan designed to test the effectiveness of Health Maintenance Organizations (HMOs).\footnote{Much of the following information is from Polich et al. (1986), particularly pp. 1-18.} The terms of the plans were 3 or 5 years. There were multiple ways in which the plans varied. We focus our attention on the variation in the coinsurance rate among plans. One plan, the free plan, provided medical care to participants without charge. We also consider plans in which 95\% of the costs of a family’s medical care was borne by the family, up to a maximum cost of $1000 or a specific percentage of the family’s income, whichever was smaller.\footnote{Using the CPI-U to adjust for inflation, the $1,000 figure would be $4,832 in 2009 dollars. For one year of the program for one subset of families, the family paid 100\% of the costs before being switched to the 95\% coinsurance plan. The RAND researchers sometimes refer to the 95\% coinsurance rate plans as the family deductible plans.} The maximum percentage of family income was one of 5, 10 or 15\%, but in our analysis we ignore these differences and group the participants in the three plans with these maximum percentages together as one treatment group. There were also three plans with 25\% coinsurance rates and three with 50\% coinsurance rates, as well as three plans with 50\% coinsurance rates for dental and outpatient mental health care and a 25\% coinsurance rates for other services; within each group of three plans, as with the 95\% plan, the maximum percentage of family income that was covered by the family could be 5\%, 10\% or 15\% for one of the three plans. In line with the RAND researchers, we group all nine 25\% and 50\% coinsurance plans together, and refer throughout the paper to this group as those on the 25\% and 50\% coinsurance plans, i.e. we ignore the variation in the maximum percentage of family income and the dental coverage rates and other variation. The maximum dollar amount that could be borne by the family for these plans varied, and was either $750 or $1000; as with the 95\% plan, the maximum amount one family could spend on medical care was the lesser of this amount and the cap on the percentage of the family’s income. The final plan was the so-called individual deductible plan, in which a family bore 95\% of outpatient medical costs, with a maximum of $150 spent per individual and $450 spent per family, but bore no share of any hospital costs. Because of the lower limits, we do not use the participants in this plan in our analysis; here we depart with the RAND researchers, who generally include the individual deductible plan with the cost-sharing plans in their analysis.\footnote{pp. 3-4, Polich et al.(1986)} Also, we do not use the HMO plan in our analysis.

In summary, in our work, we consider the group on the free plan as the treatment
group and we generally use two groups as control groups: (i) the set of all participants in any 95% coinsurance plan and (ii) the set of all participants in any of the 25%, 50% or 95% coinsurance plans.

After the baseline interview gathered demographic and self-reported health information, the overall sample for the RAND experiment was restricted to those 61 years of age or younger at enrollment, and more generally to only those that would not be eligible for Medicare at any time during the experiment. It was also restricted to the non-military, non-institutionalized population, and also military veterans with disabilities due to service were excluded. Finally, anyone whose family income was greater than $25,000 in 1973 dollars (roughly the highest 3 percent in those cities where the experiment took place) was also excluded.\textsuperscript{13}

There were six sites at which the health insurance experiment was conducted, which varied in terms of urban/rural settings and city size.\textsuperscript{14} The details of the plans did vary by sites, especially in as much as one site (Dayton, Ohio) was the site for a pilot program which comprised part of the experiment, and the researchers fine-tuned their methods after starting this program.

In selecting the initial sample to be given the baseline interview, the RAND researchers chose between 100 and 300 geographical clusters (often of about four city blocks each) in each of the six sites, and selected every $n^{th}$ dwelling unit from within the clusters.\textsuperscript{15} The families were also offered a cash incentive to participate based on their current insurance coverage, so as to ensure that they would not be hurt financially by joining the experiment. When choosing whom to offer each insurance plan, among those eligible participants who had completed the baseline interview, the experimenters used the Finite Selection Model, developed by Carl Morris (1979) particularly for use with the HIE in order to balance the characteristics of families assigned to each plan.\textsuperscript{16} Additionally, at some of the experiment sites, there was some oversampling of low-income families.\textsuperscript{17}

The demographic data and self-reported data on health is gathered largely from the baseline interview, and other interviews and questionnaires. Health data were gathered from a medical exam given at enrollment to about 60% of those on one of the experimental plans, and given at exit to 99% of those on the plans.\textsuperscript{18} Table

\textsuperscript{13}pp. 2-3, Polich et al. (1986) and p. 405, Newhouse (1993)
\textsuperscript{14}The sites were: Dayton, Ohio; Seattle; Charleston, South Carolina; Franklin County, in northwestern Massachusetts; Fitchburg and Leominster, two towns roughly 50 miles from Boston; and Georgetown County, in coastal South Carolina.
\textsuperscript{15}See p. 424, Newhouse (1993) and also the Sampling section of the Data Collection Description, given, for example, in the first few unnumbered pages of the ICPSR version of Polich et al. (1986). However, in the data files provided, we have been unable to find mention of these clusters.
\textsuperscript{16}It is similar in part to the method of Conlisk and Watts (1969), but allows for a few changes, including allowing assignment according to continuous variables in a way that one need not group such continuous variables into discrete strata. See p. 419, Newhouse (1993).
\textsuperscript{17}p. 406, Newhouse (1993).
\textsuperscript{18}p. 14, Newhouse (1993).
5.1 gives descriptions of some of the variables we use in the analysis, in particular the variables requiring more description than a short title makes clear. Much of the data available is derived data, and those preparing the data did some checking and changing of the data for consistency.\textsuperscript{19}

5.3.2 NHANES II

We also use data from the Second National Health and Nutrition Examination Survey (NHANES II), a national survey which took place in 1976 through 1980. The sampling universe consisted of the non-institutionalized, civilian population of the U.S., for people 6 months old through 74 years old.\textsuperscript{20} 20,325 people were sampled and given detailed medical examinations by a mobile unit. The elderly, preschool children and low income groups were oversampled. There is stratification into 64 strata, and one site was chosen in each stratum, and there was additional sampling by clusters within each site. The mobile examination centers were situated at each geographical site for a period of 4 to 6 weeks. At that time, researchers medically examined participants and asked questions about demographic characteristics. (Each participant was only examined once during the course of the survey; the survey lasted for multiple years as there were 64 sites but only three mobile exam centers.) Analysis weights and strata are provided but detailed lower-level clustering information is not.

5.4 Methods

5.4.1 Comparing Non-experimental Evaluation with Experimental Evaluation

We detail how we compare the estimates we obtain by comparing the effects among the RAND control and treatment groups with estimates we would obtain if we tried to use NHANES II data as a substitute for the control groups. In other words, we test how well the estimators would work if we had only a treatment group, but no randomized control group. The NHANES II data is not a very good comparison group; a glaring problem is that there is no information in the NHANES II regarding health insurance coverage. As noted above, there is reason to assume that the average person has a coinsurance rate that is certainly much higher than 0\% and thus we can at least give a comparison for the effect of free care versus average coinsurance rates (which one should interpret as an average over both insured and uninsured individuals, the latter having a coinsurance rate of 100\%). Of course, the fact that we do not have another good control group to use is, in itself, an important demonstration of how difficult it can be to arrive at a good estimate without a randomized experiment.

\textsuperscript{19}p. 18, Polich et al. (1986).
\textsuperscript{20}Information on the NHANES survey here comes from pp. 15-24 of McDowell et al. (1981).
Also, unlike the NHANES II sample, the RAND sample censors the top 3% of the income distribution, which could bias our results somewhat.\footnote{The NHANES study has information about family income, but only in ranges. The top range is too low to allow truncation of the sample by the same $25,000 cutoff as the Rand experiment.}

The two health outcomes we look at are measures of diastolic and systolic blood pressure, which are important tools used to diagnose high blood pressure: diastolic blood pressure was used as a diagnostic tool for high blood pressure in the past, whereas both are used now.\footnote{See p. 257 of Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure (1977), and Table 3, p. 1211 of Chobanian et al.} Generally speaking, lower measurements are better. We have used results on the effects on diastolic and systolic blood pressure readings largely because these health measures are available to us in the NHANES sample, but also because RAND researchers reported significant beneficial effects at the 5% level on diastolic blood pressure for those on the free plan.\footnote{See p. 201, Newhouse (1993). Other RAND researchers reported significant effects only at the 6% level for the same measure. See p. 1430, Brook et al. (1983).} Concerning the mechanism for this effect, the RAND researchers give some suggestive evidence that the improvement could be due to patients having more office visits on the free plans, during which a diagnosis of high blood pressure would have been more likely to have been made, but they allow for other possible mechanisms. It turns out that we find below that this diagnostic blood pressure result is in some ways not robust, but the comparison of the estimates is still informative. There is no such result found for systolic blood pressure, so clearly this is a much weaker test for evaluation methods to pass, seeing that, at the 5% level, 95% of the time a evaluation method that chose randomly would not find a statistically significant result.

In the RAND data, there are measurements of diastolic and systolic blood pressure, made at exit for almost all participants and at entry for those who had an enrollment medical exam (about 60% of enrollees). Generally, these were made while sitting, and if either the systolic reading was greater than or equal to 140 or the diastolic reading was greater than or equal to 90, then the measurement was made again and that second measurement was used; at most sites the right arm was used at entry and the left at exit.\footnote{There are exceptions to this; in particular, the procedure varied in other ways at entry and exit, and also at entry depending on whether or not the site was the Dayton pilot site or another site. See pp. 33-34 and 257-260, Operskalski et al. (1987).} In the NHANES II data on the other hand, each person’s blood pressure is taken while seated at the beginning of the exam and two more times at the end (once while recumbent, and again while seated). We use only the two seated blood pressures.\footnote{The mean of the first blood pressure readings is higher than the mean of the second readings, for both diastolic and systolic readings. This could be related to so-called “white coat syndrome”, which describes the situation in which blood pressure rises due to the stress created from a doctor visit; perhaps, in a similar fashion, stress subsides later in a medical exam.} The NHANES II data readings were collected according to 1951 American Heart Association guidelines, whereas there is no such detail regarding the
name of a standardized procedure used in the RAND HIE for blood pressure readings, to the best of our knowledge.\textsuperscript{26}

It is important to note that blood pressure readings can vary even “minute-to-minute” for an individual.\textsuperscript{27} So this is certainly a noisy measure to use to study health effects, and we might expect that a lot of the variability could be due to unobserved ambient effects of each person’s environment or day, and thus we might expect difficulty in detecting any significant effects in the data. Additionally, we could speculate that patients might be calmer towards the end of a medical exam, but this is a guess at best.

The RAND researchers only looked at those 14 years of age at entry and older, and the experiment was restricted to those of age at most 61 at entry.\textsuperscript{28} We generally restrict our analysis in the paper to a similar group. However, our goal is to make the control group as similar as possible to the group on the RAND free plan. In particular, we want the groups to be of similar ages at the time of the medical exam, because our key outcomes are measured at the time of the exit medical exam for the RAND survey and age is solicited for NHANES participants at the same time as their medical exams. Because many RAND plans lasted for three years, we thus restrict to those aged 17 to 64 years at exit, inclusive, and to those same ages for the NHANES survey.\textsuperscript{29} Generally speaking, in the calculations below, we use the age at exit for the RAND sample unless we specifically refer to the age at entry.

After the restriction by age, there are 10,196 people in the NHANES sample, 1,315 in the RAND free health insurance plan, 1,811 in any of the RAND 25%, 50% and 95% coinsurance rate plans, and finally 767 in the RAND 95% coinsurance rate plans.

We use the weights attached to the NHANES II data in our analysis. We also attempt to account for the oversampling of the RAND sample for low-income groups; however, weights are not given, and in fact the RAND researchers do not account for them, although they claim that this is justified by “quantitatively modest interactions

\textsuperscript{26}p.102, U.S. Department of Health and Human Services, National Center for Health Statistics (1986).

\textsuperscript{27}p. 1212, Reeves (1995).

\textsuperscript{28}There are four exceptions, but this is the adult group for which there is medical data. See p. 15 of Operskalski et al. (1987).

\textsuperscript{29}RAND participants enrolled for either 3 or 5 years, so those who enrolled for 3 years were aged 14 to 61 at the time of entry.

\textsuperscript{30}Recall that preschool children and the elderly were oversampled; this explains why our age restrictions reduce the NHANES sample from 20,325 to only 10,196. For the RAND sample, we first restrict the group of 8,254 enrollees in any plan to those who were given an adult (age 14 or older) medical exam at enrollment or exit or given an adult enrollment or exit medical history questionnaire. In total there were 5,871 such people. (See p. 15 of Operskalski et al., 1987.) The reduction from 8,254 to 5,871 is mostly explained by the restriction to those aged 14 or older. Exclusion of the HMO and individual deductible participants explain why there are only 3,126 (= 1,315 + 1,811) observations remaining in our final sample, consisting of those in the free, 25%, 50% or 95% plans.
Oversampling was determined by income cutoffs that vary according to the initial enrollment year and the family size, and our estimation of these weights is by necessity an approximation. Although there are problems with our weights, their use seem likely to present an improvement over the RAND work in which oversampling is ignored.

We briefly note how we estimate family size in our analysis. We match individuals to a family if a tag for family is available and matching them to a family head if a tag for family head is available. We take the family size to be the larger of the sizes of the two families thus defined.

Recall that usually an entire family is enrolled (or not enrolled) in a given plan in the RAND experiment. Also multiple members of the same family could be chosen for inclusion in the NHANES survey. Therefore correcting for correlation within families is important to avoid misestimation (most likely underestimation) of standard errors. Thus we correct for clustering by family group, in both the RAND and NHANES surveys.

As the data do not contain information on the four-block clusters that RAND uses or some of the more complicated clustering methods used in the NHANES survey, we do not correct for this clustering. We do have access to the clustering by site in each survey, but as there are only six sites in the RAND survey, statistical inference would be poor for that survey if we clustered by site as well; after all, the statistical inference would in that case be limited by the fact that six is clearly a small sample size for the number of cities and counties in the U.S. Thus an important assumption of this work is that the relationships studied do not vary substantially by location; this is undoubtedly incorrect, but a necessary limitation of the data. However, the point estimates, unlike the estimates of standard error and associated statistical inference, are not affected by the clustering.

We always use the RAND free insurance plan population as the treatment group. In the analysis of the evaluation methods, we treat four groups as control groups: (i) the RAND 95% plans group; (ii) the RAND 25%, 50% and 95% plans all grouped together; (iii) the NHANES group, using the first blood pressure reading (for each outcome, systolic and diastolic blood pressure); and (iv) the NHANES group, using the second blood pressure reading (again, for each outcome).

For each of the four alternative control groups, we use the following regression specifications. We regress on: (a) a dummy indicator for the free plan; (b) the free

---

32 For one, because we have access only to income data in 1973 dollars, and can not be sure about the deflator they use, we are unable to guarantee that our cutoffs are correct. There are also issues as to how exactly RAND researchers measure family size; additionally, missing income data is an issue-- if income is missing, we assume that the observation is not oversampled.
33 There are two families for which this procedure did not lead to a clear family size; inspection of the data revealed that data errors were likely the issue.
34 Thus our point estimates should be accurate, as we use the NHANES survey weights, except in as much as our estimation of weights to account for oversampling in the RAND survey is flawed.
plan dummy and a control for age; and (c) the free plan dummy and controls for age, sex, an indicator for being married, an indicator for being black, dummy variables for education (a dummy for strictly more than 8 and at most 9 grades completed, a dummy for strictly more than 9 and at most 10 grades completed, up through more than 15 and at most 16, and finally a dummy for strictly more than 16 years of education), a dummy for employment status, and finally dummies for eleven of the twelve family income categories given in the NHANES II data. In all, we have 12 regressions. Again, Table 5.1 has more detail on some of these variables; descriptions are given for definitions of the variables in the RAND data, as opposed to for their definitions in the NHANES data set. We naturally attempt to approximate the definitions as well as possible when using the NHANES data.

Now suppose we were in the hypothetical situation in which there were no randomized control group, but instead only the free health insurance plan was offered. In such a setting, a researcher may be concerned that certain characteristics could affect whether or not people enroll in the free plan offered to them under the RAND HIE. (For example, the RAND researchers report differing refusal rates by family size, and as we will see below, we find varying refusal rates by age as well.) For these reasons, we follow Lalonde and use the Heckman selection correction estimator for all combinations of the two NHANES control groups and the three sets of regressors given above. Here we would think of ourselves as being in a world with no RAND control groups, and thus the selection equation determines participation in the free plan, as opposed to staying out of the plan, where the latter is the analogue of being in the NHANES sample.

We select on family size, age, sex, working status and income. The family size variable is not included as a regression control to satisfy the exclusion restriction for the selection correction; this ensures that identification is not solely based on functional form. The family size variable, as we will see below, is higher on average for those on the free plan, and thus we might view it as being important for whether or not one selects into the free plan, and it is not as clear that it would affect blood

---

35The age variable is the age at exit for the RAND survey, as opposed to the age at enrollment, in order to best match the age from the NHANES survey, which was given at the time of the NHANES medical exam. The education, marriage and working status variables are all taken at enrollment, as these measures at exit are not available in the accessible data; this could lead to a lack of comparability with the NHANES data, in which information about these characteristics is asked at the time of the medical examination. Additionally, the employment status variable is only given for those 16 or older in the NHANES survey and we adjust the RAND data accordingly. The top education category is essentially determined by the upper education category in the NHANES survey.

36The income variable, for example, does not translate cleanly between the RAND and NHANES data, due to our need to make estimates due to different reporting years; we use the CPI to account for inflation. There is a similar issue with the married indicator, as the survey options for responding to the initial question are not the same in the NHANES II and RAND surveys— in both cases, we choose married to mean married with one’s spouse present.
pressure readings. Working status as well, for example, would seem to have an effect on selection outside of its effect on health outcomes, because of its relationship with employer-provided health insurance at the time of entry into the experiment.

5.4.2 Validity of the RAND Experiment and Findings

There are differing refusal rates for each group of plans: for those age 14 and over, 76.5% accepted the 95% coinsurance plans they were offered, 81.9% of those offered one of the 25% or 50% plans accepted, and 93.2% accepted the free plan if offered it. Because the randomization is not perfect, it is natural to question the results of the HIE. Also, the RAND researchers often use exit values predicted from a linear regression when comparing health outcomes. We find this method to be not entirely convincing; indeed, Freedman (2008) discusses a somewhat similar methodology in which analysts use multiple regression in experiments, and finds that it may be worse than a simple comparison of means. For these reasons, we are motivated to examine more carefully some of the RAND researchers’ health outcome results.

We look at some $t$-tests for comparisons of means of several measures for those on different plans, in order to give a minimal check of the observables for evidence of whether or not the assignment to plans was random. As discussed above, calculations for these comparisons account for survey weights and clustering by family. We give two-sided tests of various pre-experiment characteristics and health measures, the latter for the subgroup of participants who were given a pre-enrollment medical examination. We compare those on the free plan with those in the 25% or 50% group (we do not include the 95% plans in this analysis), and also compare those on the free plan with those in the 95% group. Recall that as mentioned above, the Rand researchers explicitly designed the randomization to balance demographic characteristics and self-reported health measures among the plans; nevertheless, differential refusal rates could affect the validity of the randomization.

After examining what are mostly demographic measures with these tests, we then focus on health measures at enrollment, using $t$-tests again to test for equality of means for the treatment and control groups.

To look more closely at the health outcome results of the HIE, we examine their result on diastolic blood pressure at exit, as well as another result RAND researchers found regarding an improvement in (exit) functional far vision. Functional far vision is measured by the best line seen on an eye chart with corrected vision (e.g., with glasses if usually worn by the participant), where the lower numbered lines are harder

---

37 There clearly may be channels by which it affects health, but we must find some variable to use for the exclusion restriction. Finding an ideal variable to satisfy the exclusion restriction is of course a disadvantage of using the selection correction to estimate experimental effects.

38 We note that refusals are not necessarily permanent—see pp.17-18 of Polich and d’Arc Taylor (1987) for details.
to see. Thus a lower number is better.\textsuperscript{39} We report estimates for coefficients of the
treatment indicators for multiple specifications. As before, for each outcome, we give
three specifications for our regressions: (i) no controls, (ii) controlling for age at the
time of the exit exam, and (iii) all of the controls given above. Despite the warnings
of Freedman (2008) against including regressors when analyzing experiments, the
practice is conventional, so we do report results with multiple regressors. We look
at both of the RAND control groups, namely (i) those in the 95% coinsurance plans
and (ii) those in the 25%, 50% or 95% coinsurance plans. (We no longer use the
NHANES control groups.) We also report results where the outcome variable is
individuals’ differences between the entry and exit values of each health measure.
These are calculated for the 60% of participants that were given an entry exam.

5.4.3 Other Outcomes of the RAND Experiment

We wish to see if there are any additional health outcomes for which there are
beneficial effects of free insurance. We first examine six key exit health outcomes,
consisting of the diastolic blood pressure and functional far vision measures for which
RAND researchers found beneficial effects, as well as four additional outcomes.\textsuperscript{40}
Two of these are functional near vision and natural far vision. Again, the functional
measures is a measure with corrective lenses; the natural is a measure without them.
We also consider an indicator for the presence of high blood pressure.

The final measure we investigate is the number of seconds it takes to walk 50 feet.
(Details on some of these variables are given in Table 5.1.) We report control and
treatment means, where we use as control groups (i) the set of all participants in any
of the 25%, 50% or 95% coinsurance plans, and (ii) the set of all participants in the
95% coinsurance plans. We also report \(p\)-values for of one-sided parametric \(t\)-tests
of the null hypothesis that treatment does not affect the mean values, against the
alternative hypothesis of a decrease in the mean value. Observe that we use one-
sided tests, as there are clear priors about the effects of free insurance on health: for
each of the six outcomes, a lower mean represents improved health.

The analysis below finds that family size seemed to be potentially related to differ-
ing relative refusal rates by plan. Thus family size might have interesting interactions
with whether or not one benefits from having free health insurance.

The same could be true of income. Free insurance might be expected to have a
larger effect on obtaining health care for a low income group than for a high income
group: the cost of health care relative to other goods and services would be lowered
much more, relative to being on a high coinsurance rate plan, for low income partic-
ipants on the free plan than for high income participants on the free plan. Thus we
may expect to in general see larger effects from being on the free plan, relative to the

\textsuperscript{39}See Table 5.1 for more description of this quantity.

\textsuperscript{40}To some extent, these examinations of the results for diastolic blood pressure and functional far
vision repeat earlier analyses; we discuss this more below.
other plans, for low income groups than for high income groups. Additionally, if low income participants had less access to medical care before the experiment than high income participants, then perhaps there are certain basic interventions provided by free health care that might have had a larger benefit to low income participants.

Thus we examine the six exit health outcomes more closely in particular for different subgroups defined by splitting along family size lines or income lines. We split family size into the category of families with 1, 2 or 3 members, and the category of families of 4 or more. We split income by the median of the household income measure.

5.5 Results

5.5.1 Comparing Non-experimental Evaluation with Experimental Evaluation

Table 5.2 presents results from regressing the outcome diastolic blood pressure, at exit, on an indicator for treatment with free insurance. For each of 12 regressions, we report estimates for the coefficient on the treatment indicator. Unlike typical reporting of regression results, we do not report estimates for coefficients of all of the covariates, but only the estimate for the coefficient on the treatment indicator— it is these coefficients we want to compare. The top number in each block is the estimate of the coefficient on the treatment indicator, with the standard error in parentheses and the \( p \)-value of a \( t \)-test for being nonzero in brackets. As mentioned above, we cluster by family.

For each row, we use a different control group. That is, in each row we pool the individuals in the RAND study who received free insurance with the individuals in the particular row’s control group. The control groups are: (i) the individuals in the RAND survey on a 95% coinsurance plan; (ii) the individuals in the RAND survey on one of the 25%, 50% or 95% coinsurance plans; (iii) the individuals in the NHANES study, using their first diastolic reading; or (iv) the individuals in the NHANES study, using their second diastolic reading. So the latter two rows are the NHANES groups that a researcher might use as a control group if a randomized controlled experiment was not available but rather one were only able to study an analogous “natural experiment”.

For each column, on the other hand, we use different sets of control variables, namely (a) no regressors, (b) age, and finally (c) all regressors, as mentioned in the above description of our methods.\(^{41}\) Observe that the estimate of the coefficient in

\(^{41}\)Recall that these are: controls for age, sex, an indicator for being married, an indicator for being black, dummy variables for grades of education completed (a dummy for more than 8, but at most 9, up through more than 15, but at most 16, and finally a dummy for strictly more than 16 years of education), a dummy for employment status, and dummies for eleven of the twelve family income
column (a) gives the straightforward estimate of the difference of the mean of the diastolic reading for the treatment group and the mean of the diastolic reading for the control group, i.e. the latter subtracted from the former.

Looking at the regression estimate represented by the first row of the first column, we see that a simple regression on a treatment indicator, within the pooled sample of those on the free plan or the 95% coinsurance rate plans, gives an estimate of \(-1.115\). This indicates an improvement in diastolic blood pressure from being on the free plan, as lower readings are generally better. As we have noted, this estimation method is exactly equivalent to computing the difference between the treatment and control groups. This estimate is statistically significant at the 5.6% level. The estimate is \(-0.773\) when using the larger group of participants in the 25%, 50% or 95% coinsurance groups, which is close to the earlier estimate and very close to the RAND estimate of \(-0.8\) which uses a control group even larger than this second group. However, our second estimate is only significant at a 10% level (see the second row). If one were to use the regression-adjusted estimates of the treatment effect, reported in the second and third columns, we would obtain fairly similar results within each row. As noted above, using regression adjustment for experiments is in general not to be advised (Freedman (2008)).

Now we turn to the NHANES “control groups”. For the initial reading of diastolic blood pressure in the doctor’s exam, we would get agreement with the RAND results in terms of the sign of the results; however, the magnitude of the estimate of the effect is much larger and we would be much more inclined to think they are statistically significant than we would with the RAND results. This is the case even if one were to use regression adjustment.

If we used the second NHANES reading on the other hand, we obtain estimates closer to the RAND estimates. However, if we chose to add multiple regressors to the analysis, as many researchers would likely do, the estimates are smaller in magnitude, and significant only at the 12% level.

In summary, the use of the NHANES observations as a control group, using a basic regression-adjustment methodology, seems to largely stand up as an estimate of the true experimental results as far as the sign of the estimate is concerned. However, the point estimates vary substantially. Also, one of the estimates, using multiple regressors and the second NHANES reading, is not statistically significant at conventional levels—so a researcher might not find a beneficial effect of the free insurance plan were they to focus on this specification.

We next turn to the systolic blood pressure reading, which RAND researchers did not appear to study for adults.\(^4\) In Table 5.3, we present results of the same kind as

---

\(^{42}\)See p. 198, Newhouse (1993). Also, recall that our largest control group does not include all of the controls that the RAND analysis does, as it excludes those on the so-called individual deductible plan.

\(^{43}\)As there were numerous articles based on the study, we can only state that it was not investigated
for diastolic blood pressure above, but now with an outcome of readings of systolic blood pressure. The systolic blood pressure is an example of a measure for which researchers should not find a result if their methodology is correct: using either set of true (RAND) control groups does not lead to any statistically significant results (at any conventional significance level), and in fact using the larger group yields a perverse point estimate.

The NHANES control groups, on the other hand, strongly point to an interpretation of a health benefit from free insurance. All of the six estimates in Table 5.3 are negative and statistically significant at the 5% level, or at much lower levels. Thus a researcher equipped with only the NHANES control groups would likely conclude that the free insurance plans had a large beneficial effect. An econometrician with access to both readings might be concerned about the large variability between the two estimates and would thus disregard the results. However, if only the first reading were available, there would be no clear reason to reject the findings; with only the second reading, it is not clear what a researcher might do. One should note, moreover, that this should be a fairly weak test of a methodology; after all, a random procedure would predict a result to be not statistically significant at the 5% level approximately 95% of the time. It is certainly a concern that this technique does not pass this weak test.

One might observe that blood pressure readings are a noisy measure, as an argument against the comparison given here. However, the standard errors are small enough that, strictly from the statistical analysis, a researcher would likely conclude that there is a beneficial effect on systolic blood pressure from free health insurance. So there is no concern of the noise from this perspective. The issue of a large variance in the measure would, however, present a strong case against comparing the RAND treatment group with the NHANES controls if one argues that there could be something distinct about the manner in which the readings were conducted in each setting. However, this is simply an argument that the control group one would use if there were no experimental control group is not sufficiently similar to the treatment group; this is of course one of the central points about why it is problematic to employ econometric techniques without an experimental control group.

We now turn to estimates using the Heckman selection correction. In Table 5.4, we report the selection correction estimates in a regression of diastolic blood pressure readings, at exit, on an indicator for being on the free insurance plan. As in Tables 5.2 and 5.3, each of the six blocks of estimates represents the estimate of this coefficient in a different specification of the main equation for the selection correction (and the associated standard error and \(p\)-value). As in the above tables, in each row, we use a different control group, while in each column, we use a different set of regressors in the main equation. In the selection equation, which is the same for all six specifications, we select on family size, age, sex, working status and eleven of the dummies for in their compendium study (Newhouse (1993)).
12 income categories. As discussed above, family size does not enter into the main equation and thus satisfies the exclusion restriction. The top number in each block is the estimate of the coefficient on the treatment indicator, with the standard error in parentheses and the \( p \)-value of a test for being nonzero in brackets. As we do throughout the paper, we adjust standard errors for clustering by family size.

The first reading of blood pressure for the NHANES survey gives an estimate fairly close to the experimental estimate of the benefit of free care, when there are no additional regressors in the main equation. It is significant at the 1% level or lower both with only a treatment indicator and when controlling for age and for multiple regressors, but is larger in magnitude for the latter two regressions. However, the latter two estimates are much larger than the experimental estimates. For the second reading, the estimate without any controls (in column 1) is very small in magnitude. While the other estimates in the row have magnitude similar to that of the experimental estimates, the estimate in the third column is statistically significant at only the 12% level. In view of the small first estimate and the lack of conventional statistical significance in the third estimate, a researcher might conclude from the second reading only that there is no evidence of a health benefit for diastolic blood pressure. So a researcher’s conclusion would be quite dependent on the reading chosen.

In Table 5.5, we report Heckman selection correction estimates for a regression of systolic blood pressure on the dummy for being on the free insurance plans. Here the first reading yields negative and statistically significant results, but the estimates adjusting for multiple regressors are much larger than the experimental estimates. Using the second reading of blood pressure in the NHANES survey and no additional regressors actually yields a positive point estimate. However, when including age as a regressor or including multiple regressors, the estimate is stable, similar to the experimental results, and close to statistically significant at conventional levels. It is not completely clear, but a researcher using the first reading alone might be inclined to conclude that the free plan had a beneficial effect on systolic blood pressure, but would view the results with caution in light of the varying results. With the second reading alone, a researcher might view the result without any regressors as a fluke or not valid due to not controlling for hidden factors; however, he or she might conclude that there is no effect. Using both the first and second readings, one might also arrive at a cautious conclusion of a benefit, but not if they were very concerned about the first estimate in the second row. Again, here, the choice of blood pressure reading would affect whether conclusions would agree with the lack of benefit found in the RAND experiment.

In summary, we have calculated the results one would obtain if one used the NHANES groups in place of the RAND control groups; we used two econometric techniques to estimate the effect on two health outcomes from being on the free insurance plan. The two techniques are: (i) using available demographic information in an effort to control for any heterogeneity in the sample; and (ii) using the Heckman selection correction to adjust for any selection bias, where in a world with no RAND
control groups one might think of this as selecting into the free plan as opposed to staying out of the free plan.

Regarding the key diastolic blood pressure result, we find that the straightforward regression methodology would yield an estimate of the same sign as the experimental estimate if one were using either NHANES reading. On the other hand, for one particular specification with multiple regressors, on which a researcher might likely place a good deal of weight when formulating conclusions, a researcher would find the result to be statistically significant at only the 12% level. Moreover, the magnitudes of some of the estimates vary widely from those in the RAND experiment. Using the selection correction for diastolic blood pressure, a researcher would find somewhat similar results, although one would find an estimate of nearly zero for one specification.

For the systolic blood pressure outcome, the results indicate the strong possibility that a researcher might find a beneficial effect of being on a free plan on systolic blood pressure, whereas the experimental results indicate that there is no clear effect. Using the Heckman selection correction with the systolic blood pressure as the outcome variable would however likely lead a researcher to make more cautious conclusions.

In summary, a researcher without access to the RAND control groups might conclude that there is a beneficial effect of the free insurance plan on diastolic blood pressure, but estimates for the magnitude effect would vary widely from the experimental estimate. And a similar researcher quite possibly could interpret the free insurance plans as having a beneficial effect on systolic blood pressure, despite the lack of any such finding in the RAND survey. The interpretation for both outcomes would depend on whether the first or second reading from the NHANES study was used.

5.5.2 Validity of the RAND Experiment and Findings

There are reasons to be concerned about the validity of the randomization and the robustness of, for example, the diastolic blood pressure result. We next address each of these issues in turn.

Examining the Randomization

As we have noted, despite offering payments designed to make participants no worse off than they would have been with their pre-experiment health insurance and to thus minimize refusals, there were nevertheless differential rates of refusal depending on the plan offered. Thus, to examine the validity of the randomization, Table 5.6 presents means, by plan assignment, of 19 demographic characteristics and health outcomes at the time of entry to the experiment. The first three columns give the means for individuals in three groups: (i) those in the free plan; (ii) those in the plans with either 25% or 50% coinsurance rates; and (iii) those in the plans with a
95% coinsurance rate. The fourth column gives \( p \)-values for a two-sided \( t \)-test of equality of means for those on the free plan vs. those on the 25% or 50% plans. The \( p \)-values for similar tests of equality of means for those on the free plan vs. those on the 95% plans are reported in the final column. (Unlike earlier analyses, we are not combining the 25%, 50% and 95% plans all together for this table, but rather we are only looking at the participants in the 25% and 50% plans.)

The average family size is larger by about .43 for those on the free plan than for those on the 25% and 50% plans. The hypothesis of equality of means is rejected for this latter comparison at a level of 2.3%. This suggests that larger families were more likely to accept the free plans than the other plans. Recall that in general entire families were enrolled, rather than individuals. Thus the result is to be expected, as the larger the family, the larger the benefit of the free plan relative to the other plans. The difference between treatment and control means is a bit smaller when comparing with those on the 95% plans, but not substantially so. As the point estimate is not much smaller, the weaker statistical significance could be due only to the smaller sample size. There is weak evidence that females were more likely to take the free plan and that those working were more likely to take the free plan; the latter is certainly unexpected.

It is certainly important to be concerned about issues of multiple testing in this setting. Two out of 19 tests in the first comparison are significant at the 5% level, and one out of 19 in the second. We focus on just the 19 tests in a particular column for simplicity (rather than grouping all 38 tests together). If each test were independent, then we would expect a 24.5% chance that at least two would be significant at the 5% level, and a 62.2% chance that at least one would be significant at the 5% level.\(^{44}\)

In Table 5.7, we turn to other health measures that one might use as outcomes of the experiment. (Table 5.1 gives more description for some of these variables.) We look at these health measures at the time of enrollment, as another test of the randomization. We again compare the participants in the free plan against two control groups– those in the 25% or 50% coinsurance rate plans and those in only the 95% coinsurance rate plans.

There is only one test that is close to statistically significant at the 5% level, namely that for natural far vision. So this table presents no clear evidence of a problem with the randomization.

Finally, we describe a simple refusal analysis by age. We regress refusal to join the RAND experiment on a constant and one's age. We do not display a table for this regression, but the coefficient of age in the regression is .00217, and it is significant at a \( p \)-value of less than .001. Thus, the older a participant is, the more likely one is to refuse the RAND enrollment offers. RAND researchers also noted this relationship and attempted to address it by controlling for age in all regressions.\(^{45}\)

\(^{44}\) Of course it is extremely unlikely that the tests are independent, and so we have likely given an overestimate of the likelihood of seeing as many significant test results as we have.

\(^{45}\) p.18, 23, Newhouse (1993).
Overall, we find some evidence to call into question the randomization setup of the experiment, especially with regard to family size and age. Although the evidence is not completely clear, it seems good to bear these concerns in mind when interpreting any of the experimental results. The RAND researchers were aware of a lower refusal rate for larger families, and they claim to account for this (except when comparing means) “either by restricting the analysis to children or by including age as an explanatory variable.”

It is surprising that under this line of reasoning, the researchers do not control directly for family size in regressions. Nevertheless, even if they were to do so, the Freedman (2008) results suggest that controlling for family size in a linear regression would not adequately adjust estimates for the differential selection.

### Examining Some of the RAND Findings

Because the researchers use exit values predicted from regressions, rather than simple means, we first examine more closely two particular findings that have been reported by RAND researchers from the experiment. We look at diastolic blood pressure and functional far vision. We first report results from regressions of functional far vision on a treatment indicator for being on the free plans relative to being on either the 95% plans, or relative to being on one of the 25%, 50% or 95% plans.

Recall that the coefficient on the treatment indicator in a regression on a constant and the treatment indicator is simply the difference of control and treatment means. Following convention, we also report regressions including other regressors. As above, each row represents a control group and each column a choice of a set of regressors, and the estimate given is for the coefficient of the treatment indicator. Note that we no longer include the NHANES control groups, as we are of course now focusing only on the RAND results.

Table 5.8 presents the results. For both sets of control groups, both with and without regressors, inclusion in the free plan appears to have improved functional vision. The estimates are statistically significant at the level of of 2% or lower. RAND researchers further examined the usage of eye exams and acquisition of corrective lenses by plan. The key reason for the difference in use of resources between plans was in increased utilization of eye exams in the free plan compared to the other plans; there was only a small difference between the plans in the percentage of participants who acquired corrective lenses once a vision problem was found.

To further test both the diastolic blood pressure and vision results, we consider the differences in the measures between entry and exit, for the 60% of the sample

---

46p. 19, Newhouse (1993)
48Analogous results for diastolic blood pressure are included in Table 5.2.
for which we have information at entry.\textsuperscript{50} As above, we present the results for two control groups and 3 sets of regressors, and the tables are presented in similar fashion. Panel A of Table 5.9 displays the results for diastolic blood pressure. We see much weaker results. Moreover, the results are of much smaller orders of magnitude than the estimates in Table 5.2 unless one includes multiple regressors.

Panel B of Table 5.9 looks at the functional far vision measure. The finding of a beneficial effect of the free plan is stronger. The straightforward difference of means, \(-.190\), for the 95\% plans control group, is significant at the 1\% level, and including a control for age, the result is similar. However, the estimate is much larger when including multiple regressors. All three estimates are of a similar magnitude, and similar to the original estimate in Table 5.8, when the 25\%, 50\% and 95\% plans are grouped together as the control group. However, in two specifications, the results are statistically significant at only the 7\% level; but this could be merely due to the sample size being 60\% of the sample size for Table 5.8.

\subsection*{5.5.3 Other Outcomes of the RAND Experiment}

We now examine more closely six health measures in the RAND study for which we have found indications that there may be benefits from being on the free plan. We look only at their values at exit, and not at differences between entry and exit. Additionally, we examine these measures for subgroups divided by family size and income, to look for differential effects in these subcategories. Income is clearly an important quantity to consider, and family size was seen to be a contributor to differential enrollment rates in Table 5.6.

Two of the health measures are discussed above, namely diastolic blood pressure and functional far vision. We include an indicator variable for high blood pressure, considering that it may be of interest, given its possible connection with diastolic blood pressure.\textsuperscript{51} We also study functional near vision and natural far vision. Finally, we include the number of seconds to walk 50 feet as a measure, having noted it as being of interest in exploratory work.\textsuperscript{52}

\textsuperscript{50}To be clear, the difference used is the entry measure subtracted from the exit measure.

\textsuperscript{51}Hypertension, or high blood pressure, is currently defined as having systolic blood pressure at least 140 or diastolic blood pressure at least 90. See Table 3, p. 1211 of Chobanian et al. In the past, it was defined in terms of diastolic blood pressure only, as mentioned earlier. See p. 257, Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure (1977). In the RAND study on the other hand, high blood pressure (HBP) was coded as 1 if the respondent had higher than normal blood pressure (140/90) or if they took medication for HBP, with the exception of the case in which the respondent had mildly elevated blood pressure (systolic of 140-159 and diastolic of at least 95) \textit{and} the condition had not previously been mentioned by a doctor. (The codebook actually does not make the definition clear for a diastolic measure of 90 to 95.) See p. 263, 266, 267 of Operskalski et al. (1987). So there should be at least a partial mechanical relationship between the high blood pressure and diastolic blood pressure variables.

\textsuperscript{52}We are not able to find the precise criterion under which the test was given, but it was given at
Table 5.10 looks at treatment and control means and differences, using (i) the participants in the 25%, 50% and 95% plans and (ii) the participants in the 95% plans as two separate control groups. The fourth column gives $p-$values for one-sided $t-$tests of the null hypothesis that the means are the same between the treatment group and control group (i) against the alternative that the treatment mean is lower (which represents better health in all six cases). We feel that the alternative hypothesis presents a clear prior, namely that free insurance should be beneficial for health; one would not \textit{a priori} expect free insurance to lead to worse health outcomes. However, there may be certainly be an argument for using two-sided tests; in this case, the reader can of course judge the results from two-sided $t-$tests by simply doubling the reported $p-$values. The final column reports $p-$values for similar (one-sided) $t-$tests, but now using the second control group. As usual, the calculations all use survey weights, and standard errors used in the $t-$tests correct for clustering by family.

A few of these tests are statistically significant at the 5% level, or close to it. For diastolic blood pressure, the $p-$value for the test for the 95% control group of .028 is exactly half that for the top left entry in Table 5.2. This is to be expected in light of two facts: (i) the coefficient estimate represented by that entry is simply an estimate of the difference between the treatment and control means; and (ii) the $p-$value there is associated with a two-sided test of the null hypothesis that the difference is zero, whereas here we are using a one-sided test.\textsuperscript{53} The functional far vision tests still show, as seen before, strong statistical evidence of benefits from free health insurance. Similarly to the blood pressure results, the $p-$values for the $t-$tests associated with these measures are exactly half (up to rounding) of the $p-$values for the regressions in the first column of Table 5.8. Of course in these cases we are not reporting new information, but we include them for completeness and also for comparison with Tables 5.11 and 5.12.

The measure of the number of seconds to walk 50 feet is improved in the free plan participants over those in both control groups, and the estimate of the difference is statistically significant at the 2% level or lower.\textsuperscript{54} Finally, we see some suggestion of effects for natural far vision, functional near vision and presence of high blood pressure, but the estimates are at most weakly statistically significant.

Because of the differential refusal rates by family size, we investigate whether any of these measures shows different levels of improvement, or lack thereof, depending on family size. We split into families of size 1, 2 or 3 for the first category, and size

\footnote{It is quite close to half for the other control group, but not exactly, as a different software package was used.}

\footnote{Because of the differential refusal rates by family size, we investigate whether any of these measures shows different levels of improvement, or lack thereof, depending on family size. We split into families of size 1, 2 or 3 for the first category, and size

\footnote{Thus it would even be significant at the 4% level using a two-sided test. The RAND researchers do not find such a result, perhaps because of the different comparison groups or perhaps because they restrict to only those individuals 35 and over who also had “joint discomfort.” See pp. 198-200 of Newhouse (1993). (Their footnotes g and h appear to have been accidentally interchanged, as a close inspection of the table and comparison to the table on pp.202-203 reveals.)}
4 or more for the second. We again consider two control groups: (i) the group of all those on plans with coinsurance rates any of 25%, 50% or 95%, and (ii) the group of all those on the 95% plans. The results are reported in Table 5.11. The first three columns report means for the control and treatment groups. As in Table 5.10, the fourth column reports $p$-values for one-sided $t$-tests of the null hypothesis of equality of means in the treatment group and (first) control group, against the alternative hypothesis that treatment improves the mean health outcome. An improvement is again indicated by a decrease in the mean, for all six of the measures. The fifth column reports $p$-values for similar tests, using the second control group. A value of NA indicates that the mean for the treatment group was actually larger than that for the respective control group, so a $p$-value for a one-sided $t$-test is not valid.

The diastolic blood pressure result is strongest for families of three or fewer people, and the measure of the seconds to walk 50 feet is strongly statistically significant, at the 0.6% level for small families. This may be due to these measures having larger impacts mostly for older individuals, whose children may already have moved away. The functional far vision measure and the functional near vision measure as well, show strong improvements for large families, on the other hand, perhaps because the transactions costs of scheduling eye exams are relatively smaller for larger families due to economies of scale.

In Table 5.12, we report the results of examining differences between treatment and control means for the 6 health outcome measures we have chosen to focus on, where we are testing the upper and lower halves of the income distribution separately. The cutoff is the median of family income in the entire RAND sample we use. This gives an income of $12,166, in 1973 dollars. Using the CPI-U, this is $58,785 in 2009 dollars.\footnote{2008 U.S. median household income was $52,029, according to the American Community Survey. (See p.4, Semega (2009).)}

Again, we report means for both the treatment group (those on the free plan) and for each of the two control groups used above. We use one-sided $t$-tests to test equality of means against the alternative hypothesis that free insurance improves the outcomes. Again, a $p$-value of NA indicates that the mean for the treatment group was actually larger than that for the respective control group, so a $p$-value for a one-sided $t$-test is not valid.

As mentioned earlier, theory would predict that we may see larger benefits from the free plan for low income groups. We do in fact see that among the lower income group there are larger improvements for the presence of high blood pressure, and the time it takes to walk 50 feet, than there are among the higher income group. By these measures, free insurance appears to have had a much larger effect in the low income population.

On the other hand, functional far vision improves much more for the higher income group, and appears to be quite a robust result-- it is statistically significant at the
1% level. Above we described how RAND researchers noted the likely importance of increased utilization of eye exams in improving this corrected vision measure. Perhaps higher income groups are more likely to take advantage of free eye exams than lower income groups. Interestingly, there is some weak evidence of an improvement in the functional near vision of the higher income group as well.

We can compare to results found by the Rand researchers. They did find a statistically significant decrease in the risk of dying for lower income high-risk individuals from being on the free plan. They also found improvements in vision and diastolic blood pressure among lower income groups that were classified as being at elevated health risk, and these improvements were larger than those among the general population. The result for vision makes our findings appear somewhat surprising, but could point to a distinction between results for elevated risk groups as opposed to those at normal risk.

5.6 Conclusion

We have considered whether econometric methods can yield estimates similar to those obtained from randomized experiments. We contribute by both offering another data point to the groundbreaking work of Lalonde (1986) on a skills training experiment, as well as by testing the applicability to experiments in other fields, by examining the RAND Health Insurance Experiment. In effect, we posit that we are in an alternate world, in which a government program has offered free insurance to a random sample of the population, but without any control groups; we then ask ourselves what estimates one would obtain from examination of what one would then call a natural experiment. We can then compare these estimates to the simple experimental estimates.

Our method is to seek a reasonable control group, namely from the NHANES survey. We consider two outcomes for which we have data available in the NHANES survey, namely readings of systolic and diastolic blood pressure. Diastolic blood pressure readings were one of the few health outcomes for which RAND researchers found a beneficial effect of free insurance. Experimental results indicate no effect on systolic blood pressure, but we include a test of this negative result as a weak test of the econometric techniques. The two techniques we examine are regression with multiple sets of regressors and the Heckman selection correction.

Focusing first on the diastolic blood pressure result, for the basic regression estimation method, we obtain results that are for the most part statistically significant at conventional levels and the same sign as the experimental estimate. However, there are two NHANES blood pressure readings for a researcher to choose from, and using

56 They use only the lower one fifth for their low income group, so we would not expect a priori for our results to be similar. See p. 197 of Newhouse (1993).
the second reading, in a regression with multiple controls, the estimate is weakly statistically significant. Moreover, the estimates of the magnitude of the treatment effects vary widely from the experimental estimate. Estimates using the Heckman selection correction also varied, and in one case give an estimate fairly close to zero. Thus a researcher’s conclusions would depend on whether or not one chose the NHANES survey exam’s first or second blood pressure reading, and on the methodology used to estimate the effect of free insurance. Thus, in using the NHANES data as a “control group” for the free RAND insurance “treatment”, an econometrician would not clearly report a positive statistically significant effect in line with the experimental estimates.

Looking at the estimation of the effect of free insurance on systolic blood pressure, there is strong statistical evidence of a beneficial effect when using the NHANES control groups and the regression adjustment method. So a researcher using only this method might conclude that there is a advantageous outcome, whereas there is no positive benefit seen from the experimental estimates. The Heckman selection correction estimates also show some very suggestive evidence for a beneficial effect of free insurance for this measure, but there are certainly also strong signs that the estimates are not robust, so a researcher may interpret the range of selection correction results as not indicative of a beneficial effect on systolic blood pressure.

One caveat regarding looking at the experimental estimates as valid estimates themselves is that there were differential refusal rates by plan that could possibly call the results into question; in particular family enrollment decisions by plan appear to have been influenced by the family’s size. On the other hand, examination of multiple health measures at enrollment gave no indication of differences by insurance plan, other than by family size.

We also examined more carefully the two key RAND results on health outcomes, namely the beneficial effects of the free plan on reducing diastolic blood pressure and improving one’s functional far vision, that is, one’s vision at distance when using corrective lenses. When examining differences between entry and exit in the diastolic blood pressure measure, among only that portion of the sample for which there are entry values, we find no statistically significant benefit to being on the free insurance plan. Overall, one should thus bear in mind that the effect on diastolic blood pressure is not absolutely clear when interpreting our results about econometric methodology.

On the other hand, we still see robust evidence for the beneficial effect on functional far vision, even when considering differences between entry and exit values for this health outcome. Also, in examining other health outcome measures, we find evidence of a beneficial effect of the free plan on an outcome unnoticed previously, to our knowledge: a measure of the number of seconds it takes to walk 50 feet, which is generally measured for those with joint problems.

We also examine several health outcomes by family size and income categories. We note some of the most interesting results for income categories. The blood pressure and walking speed results appear to be improved much more by the free plan for the
lower half of the income distribution than for the upper half. On the other hand, free insurance appears to have a stronger effect on functional vision at a distance for those in the upper half of the income distribution. We note that we did not correct these findings for multiple testing. Although the number of tests is not particularly large, this issue should be taken into consideration when interpreting the results.

One should be a bit cautious about interpreting the above results as we do not have the information to account completely for all of the clustering that occurs in the surveys. However, we do cluster at the level of the family; a seemingly clear source of correlation. Moreover, it is very doubtful that the modest changes needed to correct for all of the clustering would affect our central finding regarding the inability of econometric techniques to match the experimental estimates.

Overall, we do not find many more significant health outcomes than the RAND researchers and some of the results we do find are not completely robust. In the end, I would tend to believe that this is for two reasons: (i) many of the types of health outcomes measured could have a large behavioral component and thus do not seem to lend themselves easily to being improved by medical care; and (ii) the overall cap on deductibles as a fraction of family income were low even on the plans with high coinsurance rates (only as much as 15% at the highest, understandably for ethical reasons) – so participants in even the plans with high coinsurance rates would likely not be limited in seeking medical care if a catastrophic event occurred. These limits should be borne in mind when interpreting any results of the RAND experiment.
Table 5.1: Descriptions of Selected RAND HIE Variables.

<table>
<thead>
<tr>
<th>Demographic Characteristics and Baseline Interview Health Measures</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Married</td>
<td>Married with spouse present at the baseline interview. (Spouse absent or unknown whereabouts are also options.)</td>
</tr>
<tr>
<td>Years of Education</td>
<td>Years of school completed (for adults) at baseline. Top code (22, 24 or none) varied by site.</td>
</tr>
<tr>
<td>In College</td>
<td>Attending college at baseline interview.</td>
</tr>
<tr>
<td>Income has the term In College</td>
<td>Income family income in the year before enrollment, in 1973 dollars.</td>
</tr>
<tr>
<td>AFDC</td>
<td>Received assistance from Aid to Families with Dependent Children at baseline.</td>
</tr>
<tr>
<td>Insured</td>
<td>Whether or not at least one “family member had health insurance of any kind” at the baseline interview.</td>
</tr>
<tr>
<td>Pre-Experiment</td>
<td>Self-reported worry about health in past year, at baseline. Scale of 1 to 4, where 1 is a great deal and 4 is none.</td>
</tr>
<tr>
<td>Worry about Health</td>
<td>Self-reported pain frequency in past year, at baseline. Scale of 1 to 4, where 1 is very often, and 4 is not at all.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Health Measures, at Entry (Enrollment) or Exit</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest Pain Indicator</td>
<td>Self-reported presence of “pain, discomfort or pressure” in chest in last year.</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>Blood cholesterol measurement, in mg/dl; second test taken if first abnormal.</td>
</tr>
<tr>
<td>On Cholesterol Diet</td>
<td>Indicator for being on a low fat or low cholesterol diet.</td>
</tr>
<tr>
<td>Functional Far Vision</td>
<td>Test of far “vision with corrective lenses, if normally worn”; denotes the best line on eye chart seen with better eye, where lower-numbered lines indicate better vision. For example, being able to see line 1 represents 20/15 vision, to see at best line 2 represents 20/20 vision; this continues (in irregular steps) to line 9, which represents 20/100 vision, and finally to line 10, which represents worse than 20/100 vision.</td>
</tr>
<tr>
<td>Functional Near Vision</td>
<td>Similar to above but for near vision.</td>
</tr>
<tr>
<td>Natural Far Vision</td>
<td>Measure of best far vision without corrective lenses, range 13 to 800, representing 13/20 to 800/20.</td>
</tr>
<tr>
<td>Natural Near Vision</td>
<td>Measure of best near vision without corrective lenses, range 15 to 400, representing 15/20 to 400/20.</td>
</tr>
<tr>
<td>Acne Indicator</td>
<td>Presence of acne, as determined from a photograph.</td>
</tr>
<tr>
<td>Acne Severity</td>
<td>Severity of acne, scale of 0 to 4.</td>
</tr>
<tr>
<td>Acne Impact</td>
<td>Indicator for some or no impact of acne.</td>
</tr>
<tr>
<td>Angina</td>
<td>Medically diagnosed presence of Angina Pectoris (chest pain often associated with coronary heart disease).</td>
</tr>
<tr>
<td>Smoking Status</td>
<td>Categorical variable for having smoked cigarettes: 1 for never, 2 for ex-smoker 3 for current smoker.</td>
</tr>
<tr>
<td>COAD</td>
<td>Presence of Chronic Obstructive Airway Disease.</td>
</tr>
<tr>
<td>Hearing Loss</td>
<td>Binary indicator for hearing loss in the better ear.</td>
</tr>
<tr>
<td>HBP Impact</td>
<td>Indicator for whether high blood pressure “caused any pain, worry/concern, or activity restrictions.”</td>
</tr>
<tr>
<td>Seconds To Walk</td>
<td>Number of seconds that it takes to walk 50 feet, which can be assisted with a cane or another person’s help if needed. Asked mostly for those with joint problems.</td>
</tr>
<tr>
<td>Grip Strength</td>
<td>Best of 3 tries with the weaker hand, in mm Hg.</td>
</tr>
<tr>
<td>UTI Indicator</td>
<td>Indicator for urinary tract infection, asked among females only. 1 if positive urine culture test, but also if “taking prescribed medications” for UTI even if a negative test. Also, following Rand data, 1 for a single pregnant female observation, regardless of test result.</td>
</tr>
<tr>
<td>UTI Impact</td>
<td>Indicator of whether urinary tract infection caused “pain, worry, concern, or curtailment of activities during the past 3 months or kept the participant bedridden in the past 30 days.”</td>
</tr>
<tr>
<td>Collapsed Ulcer</td>
<td>1 for an active ulcer; 2 dyspepsia; 3 an old, cured or silent ulcer; 4 no disease.</td>
</tr>
</tbody>
</table>

Notes: Quoted text is taken from Polich et al. (1986) or Operskalski et al. (1987).
Table 5.2: Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Diastolic Blood Pressure on Treatment Indicator and Various Controls

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Set of Regressors</th>
<th>Estimate for Treatment Indicator (Std. Error) [p-value]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dummy only</td>
<td>Age</td>
</tr>
<tr>
<td>RAND Controls (95)</td>
<td></td>
<td>-1.115 [(0.58)[0.056]]</td>
</tr>
<tr>
<td>RAND Controls (25/50/95)</td>
<td></td>
<td>-0.773 [(0.48)[0.1]]</td>
</tr>
<tr>
<td>NHANES (Diastolic 1)</td>
<td></td>
<td>-2.413 [(0.39)[0]]</td>
</tr>
<tr>
<td>NHANES (Diastolic 2)</td>
<td></td>
<td>-1.335 [(0.39)[0.001]]</td>
</tr>
</tbody>
</table>

Notes: Estimates of the coefficient on a treatment indicator for 12 distinct regressions are displayed in the table, one for each combination of 4 choices of control groups and 3 choices of sets of regressors. Standard errors are in parentheses and p-values are in brackets. A p-value of 0 indicates that \( p < 0.0005 \) holds. The 4 choices of control groups, which vary along the rows, are: (i) the RAND controls with 95% coinsurance; (ii) the RAND controls with 25, 50 and 95% coinsurance, all grouped together; (iii) the first NHANES reading of diastolic blood pressure, denoted by Diastolic 1; and (iv) the second NHANES reading of diastolic blood pressure, denoted by Diastolic 2. The 3 choices of sets of regressors, which vary along the columns, are: (i) only a dummy treatment indicator; (ii) a dummy treatment indicator and an age regressor; and (iii) all regressors, as described in the text.
Table 5.3: Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Systolic Blood Pressure on Treatment Indicator and Various Controls

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Dummy only</th>
<th>Age</th>
<th>All Regressors</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAND Controls (95)</td>
<td>−0.635 (0.93)[0.49]</td>
<td>−0.638 (0.84)[0.45]</td>
<td>−1.076 (0.94)[0.25]</td>
</tr>
<tr>
<td>RAND Controls (25/50/95)</td>
<td>0.335 (0.71)[0.64]</td>
<td>0.371 (0.64)[0.56]</td>
<td>−0.209 (0.71)[0.77]</td>
</tr>
<tr>
<td>NHANES (Systolic 1)</td>
<td>−2.873 (0.56)[0]</td>
<td>−2.458 (0.5)[0]</td>
<td>−3.133 (0.57)[0]</td>
</tr>
<tr>
<td>NHANES (Systolic 2)</td>
<td>−1.354 (0.56)[0.015]</td>
<td>−1.012 (0.5)[0.045]</td>
<td>−1.216 (0.57)[0.033]</td>
</tr>
</tbody>
</table>

Notes: Estimates of the coefficient on a treatment indicator for 12 distinct regressions are displayed in the table, one for each combination of 4 choices of control groups and 3 choices of sets of regressors. Standard errors are in parentheses and p-values are in brackets. A p-value of 0 indicates that $p < 0.0005$ holds. The 4 choices of control groups, which vary along the rows, are: (i) the RAND controls with 95% coinsurance; (ii) the RAND controls with 25, 50 and 95% coinsurance, all grouped together; (iii) the first NHANES reading of systolic blood pressure, denoted by Systolic 1; and (iv) the second NHANES reading of systolic blood pressure, denoted by Systolic 2. The 3 choices of sets of regressors, which vary along the columns, are: (i) only a dummy treatment indicator; (ii) a dummy treatment indicator and an age regressor; and (iii) all regressors, as described in the text.
Table 5.4: Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Diastolic Blood Pressure on Treatment Indicator and Different Sets of Controls, using Heckman Selection Correction.

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Set of Regressors</th>
<th>Estimate for Treatment Indicator (Std. Error)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dummy only</td>
<td>Age</td>
<td>All Regressors</td>
</tr>
<tr>
<td>NHANES (Diastolic 1)</td>
<td>-1.098 (0.42)</td>
<td>-1.873 (0.4)</td>
<td>-1.839 (0.4)</td>
</tr>
<tr>
<td>NHANES (Diastolic 2)</td>
<td>-0.0373 (0.42)</td>
<td>-0.728 (0.4)</td>
<td>-0.62 (0.4)</td>
</tr>
</tbody>
</table>

Notes: Estimates of the coefficient on a treatment indicator for 6 distinct Heckman selection correction models are displayed in the table, one for each combination of 2 choices of control groups and 3 choices of sets of regressors. Standard errors are in parentheses and p-values are in brackets. A p-value of 0 indicates that $p < 0.0005$ holds. The 2 choices of control groups, which vary along the rows, are: (i) the first NHANES reading of diastolic blood pressure, denoted by Diastolic 1; and (ii) the second NHANES reading of diastolic blood pressure, denoted by Diastolic 2. The 3 choices of sets of regressors, which vary along the columns, are: (i) only a dummy treatment indicator; (ii) a dummy treatment indicator and an age regressor; and (iii) all regressors, as described in the text. The selection equation uses age, sex, working status, family income dummy variables and family size; the family size variable is not included in the main equation.
Table 5.5: Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Systolic Blood Pressure on Treatment Indicator and Different Sets of Controls, using Heckman Selection Correction

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Set of Regressors</th>
<th>Dummy only</th>
<th>Age</th>
<th>All Regressors</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHANES (Systolic 1)</td>
<td>Estimate for Treatment Indicator (Std. Error) [p-value]</td>
<td>-1.524 (0.63)[0.016]</td>
<td>-2.898 (0.55)[0]</td>
<td>-3.131 (0.57)[0]</td>
</tr>
<tr>
<td>NHANES (Systolic 2)</td>
<td></td>
<td>0.185 (0.61)[0.76]</td>
<td>-1.047 (0.55)[0.058]</td>
<td>-1.216 (0.57)[0.033]</td>
</tr>
</tbody>
</table>

Notes: Estimates of the coefficient on a treatment indicator for 6 distinct Heckman selection correction models are displayed in the table, one for each combination of 2 choices of control groups and 3 choices of sets of regressors. Standard errors are in parentheses and p-values are in brackets. A p-value of 0 indicates that $p < 0.0005$ holds. The 2 choices of control groups, which vary along the rows, are: (i) the first NHANES reading of systolic blood pressure, denoted by Systolic 1; and (ii) the second NHANES reading of systolic blood pressure, denoted by Systolic 2. The 3 choices of sets of regressors, which vary along the columns, are: (i) only a dummy treatment indicator; (ii) a dummy treatment indicator and an age regressor; and (iii) all regressors, as described in the text. The selection equation uses age, sex, working status, family income dummy variables and family size; the family size variable is not included in the main equation.
Table 5.6: Comparison of Means of Various Characteristics and Health Measures at Entry, and Parametric Two-Sided $t$–tests of Equality of Means.

<table>
<thead>
<tr>
<th></th>
<th>Free Plan</th>
<th>Means</th>
<th>$p$–value of $t$–test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>25/50%</td>
<td>95%</td>
</tr>
<tr>
<td>Quantity</td>
<td></td>
<td>Coins.</td>
<td>Coins.</td>
</tr>
<tr>
<td>Age at Enrollment</td>
<td>32.53</td>
<td>32.96</td>
<td>32.17</td>
</tr>
<tr>
<td>Female</td>
<td>0.521</td>
<td>0.526</td>
<td>0.551</td>
</tr>
<tr>
<td>Black</td>
<td>0.125</td>
<td>0.111</td>
<td>0.145</td>
</tr>
<tr>
<td>Married</td>
<td>0.647</td>
<td>0.67</td>
<td>0.633</td>
</tr>
<tr>
<td>Family Size</td>
<td>4.614</td>
<td>4.187</td>
<td>4.253</td>
</tr>
<tr>
<td>Years of Education</td>
<td>11.92</td>
<td>12.16</td>
<td>12.2</td>
</tr>
<tr>
<td>Attending College</td>
<td>0.0373</td>
<td>0.044</td>
<td>0.0463</td>
</tr>
<tr>
<td>Working</td>
<td>0.64</td>
<td>0.68</td>
<td>0.649</td>
</tr>
<tr>
<td>Full Time</td>
<td>0.854</td>
<td>0.838</td>
<td>0.811</td>
</tr>
<tr>
<td>Income (1973 $)</td>
<td>11560</td>
<td>12030</td>
<td>11730</td>
</tr>
<tr>
<td>AFDC</td>
<td>0.0433</td>
<td>0.0304</td>
<td>0.0344</td>
</tr>
<tr>
<td>Insured Pre-Experiment</td>
<td>0.882</td>
<td>0.904</td>
<td>0.908</td>
</tr>
<tr>
<td>Worry About Health</td>
<td>3.296</td>
<td>3.298</td>
<td>3.316</td>
</tr>
<tr>
<td>Pain Frequency</td>
<td>3.203</td>
<td>3.223</td>
<td>3.186</td>
</tr>
<tr>
<td>Diastolic</td>
<td>74.42</td>
<td>74.13</td>
<td>74.59</td>
</tr>
<tr>
<td>Systolic</td>
<td>123.3</td>
<td>123</td>
<td>121.9</td>
</tr>
<tr>
<td>Chest Pain Indicator</td>
<td>0.163</td>
<td>0.156</td>
<td>0.153</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>202.1</td>
<td>205.6</td>
<td>206.8</td>
</tr>
<tr>
<td>Functional Far Vision</td>
<td>2.319</td>
<td>2.4</td>
<td>2.281</td>
</tr>
</tbody>
</table>

Notes: See text for details and Table 5.1 for further descriptions of many of the variables.
Table 5.7: Comparison of Means of Additional Health Measures at Entry, and Parametric Two-Sided \( t \)-tests of Equality of Means.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Free Plan</th>
<th>Means</th>
<th>( p )-value of ( t )-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>25/50%</td>
<td>95%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coins.</td>
<td>Coins. Rate</td>
</tr>
<tr>
<td>Acne Indicator</td>
<td>0.174</td>
<td>0.175</td>
<td>0.153</td>
</tr>
<tr>
<td>Acne Severity</td>
<td>1.912</td>
<td>2.098</td>
<td>1.791</td>
</tr>
<tr>
<td>Acne Impact</td>
<td>0.165</td>
<td>0.173</td>
<td>0.163</td>
</tr>
<tr>
<td>Angina</td>
<td>0.00107</td>
<td>0.0142</td>
<td>0.0111</td>
</tr>
<tr>
<td>Smoking Status</td>
<td>1.904</td>
<td>1.925</td>
<td>1.907</td>
</tr>
<tr>
<td>COAD</td>
<td>0.0975</td>
<td>0.101</td>
<td>0.0969</td>
</tr>
<tr>
<td>Hearing Loss</td>
<td>0.0249</td>
<td>0.0472</td>
<td>0.0319</td>
</tr>
<tr>
<td>On Cholesterol Diet</td>
<td>0.308</td>
<td>0.28</td>
<td>0.391</td>
</tr>
<tr>
<td>High Blood Pressure</td>
<td>0.12</td>
<td>0.126</td>
<td>0.106</td>
</tr>
<tr>
<td>HBP Medication Use</td>
<td>0.477</td>
<td>0.514</td>
<td>0.475</td>
</tr>
<tr>
<td>HBP Impact</td>
<td>0.47</td>
<td>0.386</td>
<td>0.446</td>
</tr>
<tr>
<td>Seconds To Walk 50 ft</td>
<td>8.772</td>
<td>8.744</td>
<td>8.74</td>
</tr>
<tr>
<td>Grip Strength</td>
<td>239.3</td>
<td>240.9</td>
<td>238.4</td>
</tr>
<tr>
<td>UTI Indicator</td>
<td>0.0766</td>
<td>0.0962</td>
<td>0.0549</td>
</tr>
<tr>
<td>UTI Impact</td>
<td>0.229</td>
<td>0.205</td>
<td>0.246</td>
</tr>
<tr>
<td>Collapsed Ulcer</td>
<td>3.623</td>
<td>3.627</td>
<td>3.551</td>
</tr>
<tr>
<td>Natural Far Vision</td>
<td>55.01</td>
<td>58.43</td>
<td>64.36</td>
</tr>
<tr>
<td>Natural Near Vision</td>
<td>73.41</td>
<td>71.57</td>
<td>65.31</td>
</tr>
<tr>
<td>Functional Near Vision</td>
<td>1.656</td>
<td>1.673</td>
<td>1.579</td>
</tr>
</tbody>
</table>

Notes: See text for details and Table 5.1 for further descriptions of many of the variables.
Table 5.8: Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Functional Far Vision Measure at Exit on Treatment Indicator and Various Sets of Regressors, using Two Control Groups

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Set of Regressors</th>
<th>Estimate for Treatment Indicator (Std. Error) [p-value]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dummy only</td>
<td>Age</td>
</tr>
<tr>
<td>RAND Controls (95)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rand Controls (25/50/95)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-.14 (.061)[.021]</td>
<td>-.14 (.06)[.019]</td>
</tr>
<tr>
<td></td>
<td>-.109 (.0487)[.022]</td>
<td>-.108 (.046)[.02]</td>
</tr>
</tbody>
</table>

Notes: Estimates of the coefficient on a treatment indicator for 6 distinct regressions are displayed in the table, one for each combination of 2 choices of control groups and 3 choices of sets of regressors. Standard errors are in parentheses and p-values are in brackets. The 2 choices of control groups, which vary along the rows, are: (i) the RAND controls with 95% coinsurance; and (ii) the RAND controls with 25, 50 and 95% coinsurance, all grouped together. The 3 choices of sets of regressors, which vary along the columns, are: (i) only a dummy treatment indicator; (ii) a dummy treatment indicator and an age regressor; and (iii) all regressors, as described in the text. See Table 5.1 for definition of outcome variable.
Table 5.9: Comparison of Estimates of Coefficient of Treatment Indicator, in Regressions of Differences in Outcome Measures between Entry and Exit, on Treatment Indicator and Various Sets of Regressors, using two Control Groups

Panel A. Outcome: Difference in Diastolic Blood Pressure from Entry to Exit

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Estimate for Treatment Indicator (Std. Error)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dummy only</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND Controls (95)</td>
<td>-.319 (.897)[.72]</td>
<td>[.72]</td>
</tr>
<tr>
<td>RAND Controls (25/50/95)</td>
<td>-.693 (.658)[.29]</td>
<td>[.30]</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND Controls (95)</td>
<td>-.319 (.896)[.72]</td>
<td>[.72]</td>
</tr>
<tr>
<td>RAND Controls (25/50/95)</td>
<td>-.689 (.658)[.30]</td>
<td>[.30]</td>
</tr>
<tr>
<td><strong>All Regressors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND Controls (95)</td>
<td>-.842 (.968)[.39]</td>
<td>[.39]</td>
</tr>
<tr>
<td>RAND Controls (25/50/95)</td>
<td>-.989 (.688)[.15]</td>
<td>[.15]</td>
</tr>
</tbody>
</table>

Panel B. Outcome: Difference in Functional Far Vision from Entry to Exit

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Estimate for Treatment Indicator (Std. Error)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dummy only</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND Controls (95)</td>
<td>-.190 (.074)[.010]</td>
<td>[.010]</td>
</tr>
<tr>
<td>RAND Controls (25/50/95)</td>
<td>-.113 (.062)[.067]</td>
<td>[.067]</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND Controls (95)</td>
<td>-.189 (.073)[.009]</td>
<td>[.009]</td>
</tr>
<tr>
<td>RAND Controls (25/50/95)</td>
<td>-.11 (.061)[.072]</td>
<td>[.072]</td>
</tr>
<tr>
<td><strong>All Regressors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND Controls (95)</td>
<td>-.277 (.08)[.0006]</td>
<td>[.0006]</td>
</tr>
<tr>
<td>RAND Controls (25/50/95)</td>
<td>-.133 (.065)[.043]</td>
<td>[.043]</td>
</tr>
</tbody>
</table>

Notes: Each panel of the table displays estimates of the coefficient on a treatment indicator for 6 distinct regressions, one for each combination of 2 choices of control groups and 3 choices of sets of regressors. See text for details on the 2 choices of control groups, which vary along the rows, and the 3 choices of sets of regressors, which vary along the columns. Standard errors are in parentheses and p-values are in brackets. The outcome variable for Panel A is the difference in diastolic blood pressure from entry to exit, while for Panel B it is the difference in functional far vision from entry to exit. See text and Table 5.1 for definitions of the diastolic blood pressure and functional far vision variables, respectively.
Table 5.10: Comparison of Mean Outcomes at Exit, and Parametric One-Sided \( t \)-tests of Lack of Change

<table>
<thead>
<tr>
<th>Health Outcome</th>
<th>Means</th>
<th>( p )-value of ( t )-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control Groups</td>
<td>Treatment Free Plan</td>
</tr>
<tr>
<td></td>
<td>25/50 and 95% Coins. Rate</td>
<td>Free Plan</td>
</tr>
<tr>
<td>Diastolic</td>
<td>78.39 78.73</td>
<td>77.61</td>
</tr>
<tr>
<td>High Blood Pressure</td>
<td>0.175 0.188</td>
<td>0.161</td>
</tr>
<tr>
<td>Seconds to Walk 50 Feet</td>
<td>9.753 9.835</td>
<td>9.358</td>
</tr>
<tr>
<td>Natural Far Vision</td>
<td>68.22 71.28</td>
<td>63.56</td>
</tr>
<tr>
<td>Functional Far Vision</td>
<td>2.339 2.371</td>
<td>2.231</td>
</tr>
<tr>
<td>Functional Near Vision</td>
<td>1.527 1.53</td>
<td>1.462</td>
</tr>
</tbody>
</table>

Notes: See text for details, and text and Table 5.1 for descriptions of variables.
Table 5.11: Comparison of Mean Outcomes at Exit, and Parametric One-Sided $t$–tests of Lack of Change, Within Subpopulations Categorized by Family Size

<table>
<thead>
<tr>
<th>Health Outcome</th>
<th>Control Groups</th>
<th>Treatment</th>
<th>$p$–value of $t$–test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Means</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25/50 and 95% Coins. Rate</td>
<td>Free Plan</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diastolic (Family Size 1 to 3)</td>
<td>80.9</td>
<td>79.41</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>81.21</td>
<td></td>
<td>0.036</td>
</tr>
<tr>
<td>Diastolic (Family Size ≥ 4)</td>
<td>76.79</td>
<td>76.64</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>77.23</td>
<td></td>
<td>0.19</td>
</tr>
<tr>
<td>High Blood Pressure (Family Size 1 to 3)</td>
<td>0.262</td>
<td>0.257</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>0.274</td>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td>High Blood Pressure (Family Size ≥ 4)</td>
<td>0.12</td>
<td>0.108</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>0.136</td>
<td></td>
<td>0.098</td>
</tr>
<tr>
<td>Seconds to Walk 50 Feet (Family Size 1 to 3)</td>
<td>10.19</td>
<td>9.508</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>10.39</td>
<td></td>
<td>0.0063</td>
</tr>
<tr>
<td>Seconds to Walk 50 Feet (Family Size ≥ 4)</td>
<td>9.394</td>
<td>9.231</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>9.429</td>
<td></td>
<td>0.26</td>
</tr>
<tr>
<td>Natural Far Vision (Family Size 1 to 3)</td>
<td>79.75</td>
<td>77.97</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>90.81</td>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td>Natural Far Vision (Family Size ≥ 4)</td>
<td>60.92</td>
<td>55.75</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>59.74</td>
<td></td>
<td>0.26</td>
</tr>
<tr>
<td>Functional Far Vision (Family Size 1 to 3)</td>
<td>2.397</td>
<td>2.332</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>2.477</td>
<td></td>
<td>0.066</td>
</tr>
<tr>
<td>Functional Far Vision (Family Size ≥ 4)</td>
<td>2.303</td>
<td>2.176</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>2.308</td>
<td></td>
<td>0.045</td>
</tr>
<tr>
<td>Functional Near Vision (Family Size 1 to 3)</td>
<td>1.613</td>
<td>1.597</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>1.551</td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Functional Near Vision (Family Size ≥ 4)</td>
<td>1.472</td>
<td>1.389</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>1.518</td>
<td></td>
<td>0.024</td>
</tr>
</tbody>
</table>

Notes: See text for details, and text and Table 5.1 for descriptions of variables.
Table 5.12: Comparison of Mean Outcomes at Exit, and Parametric One-Sided \( t \)-tests of Lack of Change, Within Subpopulations Categorized by Income

<table>
<thead>
<tr>
<th>Health Outcome</th>
<th>Means</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control Groups</td>
<td>Treatment Free Plan</td>
</tr>
<tr>
<td></td>
<td>25/50 and 95% Coins. Rate</td>
<td>Free vs. 25/50 and 95%</td>
</tr>
<tr>
<td>Diastolic (Lower Income Half)</td>
<td>78.36 78.71 77.54</td>
<td>0.1 0.081</td>
</tr>
<tr>
<td>Diastolic (Upper Income Half)</td>
<td>78.23 78.7 77.36</td>
<td>0.12 0.061</td>
</tr>
<tr>
<td>High Blood Pressure (Lower Income Half)</td>
<td>0.167 0.181 0.135</td>
<td>0.053 0.042</td>
</tr>
<tr>
<td>High Blood Pressure (Upper Income Half)</td>
<td>0.174 0.193 0.187</td>
<td>NA 0.43</td>
</tr>
<tr>
<td>Seconds to Walk 50 Feet (Lower Income Half)</td>
<td>9.956 10.1 9.317</td>
<td>0.0071 0.014</td>
</tr>
<tr>
<td>Seconds to Walk 50 Feet (Upper Income Half)</td>
<td>9.587 9.594 9.36</td>
<td>0.18 0.24</td>
</tr>
<tr>
<td>Natural Far Vision (Lower Income Half)</td>
<td>66.95 78.23 63.6</td>
<td>0.28 0.045</td>
</tr>
<tr>
<td>Natural Far Vision (Upper Income Half)</td>
<td>71.03 67.46 62.89</td>
<td>0.092 0.27</td>
</tr>
<tr>
<td>Functional Far Vision (Lower Income Half)</td>
<td>2.358 2.394 2.294</td>
<td>0.17 0.13</td>
</tr>
<tr>
<td>Functional Far Vision (Upper Income Half)</td>
<td>2.304 2.301 2.116</td>
<td>0.0021 0.011</td>
</tr>
<tr>
<td>Functional Near Vision (Lower Income Half)</td>
<td>1.557 1.565 1.502</td>
<td>0.17 0.22</td>
</tr>
<tr>
<td>Functional Near Vision (Upper Income Half)</td>
<td>1.464 1.456 1.371</td>
<td>0.034 0.086</td>
</tr>
</tbody>
</table>

Notes: See text and Table 5.1 for descriptions of variables, and text for definition of lower and upper halves of the income distribution, as well as more detail generally.
Bibliography


