Performance Pay and Wage Inequality*

Thomas Lemieux

W. Bentley MacLeod

Daniel Parent

Abstract

An increasing fraction of jobs in the U.S. labor market explicitly pay workers for their performance using bonus pay, commissions, or piece-rate contracts. Using data from the Panel Study of Income Dynamics, we show that compensation in performance-pay jobs is more closely tied to both observed and unobserved productive characteristics of workers than in non-performance-pay jobs. We also find that the return to these productive characteristics increased faster over time in performance-pay than non-performance-pay jobs. We show that this finding is consistent with the view that underlying changes in returns to skill due, for instance, to technological change, induce more firms to offer performance-pay contracts, and result in more wage inequality among workers who are paid for performance. Thus, performance pay provides a channel through which underlying changes in returns to skill get translated into higher wage inequality. We conclude that this channel accounts for 21 percent of the growth in the variance of male wages between the late 1970s and the early 1990s, and for most of the increase in wage inequality above the 80th percentile over the same period.

*The authors would like to thank David Card, Lawrence Katz, David Lee, Edward Lawler, Edward Leamer, Edward Lazear, Emmanuel Saez and participants at the IZA Workshop on Labor Market Institutions, NBER Labor Studies Workshop and Cornell University for helpful discussions, and five anonymous referees for their comments. We also thank Dan Carvell for very able research assistance.

I. Introduction

The standard competitive model of the labor market supposes that wages are equal to marginal products and that the wage structure is determined by the equilibrium of supply and demand. That simple model forms the backbone of most studies of the evolution of wage inequality. For example, Katz and Murphy (1992) argue that the return to education increased in the 1980s because the rate of increase in the relative supply of more-educated labor decelerated, while relative demand steadily increased. Similarly, Juhn, Murphy, and Pierce (1993) argue that the growth in within-group wage inequality throughout the 1970s and 1980s was driven by an increase in the demand for unobserved skills. A main virtue of such studies that use a standard competitive model of the labor market is that they generally provide a straightforward interpretation of the evolution of the wage structure in familiar terms of the supply and demand for different types of labor.

Despite the appeal of the standard competitive model, it is also well established that it is, at best, only a good approximation for the way wages are actually set in the labor market. In particular, when markets are imperfect and information is costly, wages are not generally equal to the productivity of workers. As a result of those frictions, the distribution of wages does not always accurately represent the distribution of workers' productivities. But as long as the legal, institutional, and contractual arrangements that determine the relationship between wages and productivity remain constant over time, the competitive model will still provide an accurate account of the changes in the distribution of wages. Whether or not this is the case is crucial to our understanding of why wage inequality increased so much over the last thirty years.

In this paper, we use data from the Panel Study of Income Dynamics (PSID) to investigate how one particular form of contractual arrangements, performance-pay schemes, have contributed to changes in wage inequality. Since the intent of these schemes is to more closely align the wage with productivity, our empirical strategy builds upon the idea that the wages of performance-pay workers are more closely linked to productive ability than are the wages of non-performance-pay workers. This can result in increasing wage inequality as the fraction of workers being paid for performance grows over time, or as the inequality enhancing effect of performance pay grows because of other underlying changes in the return to productive ability.

There are several reasons why studying the link between performance pay and wage inequality is particularly appealing. First, there has been a steep growth (in our PSID data and in other data sources) in the fraction of workers who are paid for performance, which suggests that these two phenomena may be closely linked. Second, performance-pay workers tend to be concentrated in the upper end of the wage distribution, which is precisely where wage inequality has grown the most dramatically over time (see, for example, Piketty and Saez (2003) and Autor, Katz, and Kearney (2006)). Third, it is well known that among executives at the very top end of the wage distribution, performance pay (bonuses, stock options, etc.) accounts for the lion's share of the growth in the level of compensation (see, e.g., Piketty and Saez (2003)), and for much of the dispersion in compensation (Frydman and Saks (2007)) in this segment of the labor market. Not surprisingly, performance pay also accounts for most of the *growth* in inequality among top executives.[1] Equipped with our PSID data, we can investigate whether this phenomenon extends to a broader cross-section of the workforce.

While we would ultimately want to know whether the growth in the use of performance pay is one of the underlying causal factors behind the growth in wage inequality, answering this question raises a number of difficult conceptual and measurement challenges. On the conceptual side, the key question is whether the growth in performance pay is driven by a set of exogenous factors unrelated to other aspects of the labor market, or is instead a rational response by firms to the same underlying factors responsible for the growth in wage inequality. For sure, measuring and rewarding individual performance is difficult and costly (see Bishop (1987)). One possible view is that performance pay has become more prevalent because the cost of collecting and processing information has declined over time with advances in information and communication technologies. Under this interpretation, one could view the growth in performance pay as a causal factor behind at least some of the growth in wage inequality.

An equally plausible alternative scenario is that as the demand for highly productive workers increases, the benefit of implementing a performance pay system outweighs the costs of introducing new measurement instruments. Under this alternative view, factors such as technological change and globalization that increase the relative demand for highly

[1] We used the ExecuComp dataset to look at changes in compensation inequality between 1992 and 2005. We found that the standard deviation (or the 90-10 gap) in log base pay has hardly changed over time, going from 0.61 in 1992 to 0.62 in 2005. By contrast, the standard deviation in the log of a broader measure of compensation (base salary and bonuses) increased from 0.72 to 0.85 over the same period.

productive workers are the underlying causal factors behind the growth in both wage inequality and the prevalence of performance-pay schemes. Even under this alternative scenario, performance pay remains a key channel through which underlying changes in the supply and demand for different groups of workers get translated into a widening wage distribution.

On the measurement side, a major challenge is to classify workers who are paid for performance, and those who are not. Unfortunately, large representative surveys like the Current Population Survey (CPS) do not contain questions that can be used to identify performance-pay jobs. Even in the PSID, all we know is whether a given worker in a given year received some pay in the form of bonuses, commissions, or piece rates. While these forms of payment correspond to the way performance pay is usually implemented in practice, we may not observe any of these in some years for workers on bonus pay if they did not merit a bonus in those years. Fortunately, the longitudinal nature of the PSID data enables us to look at whether a worker *ever* received bonuses, commissions, or piece rates on his or her current job, which provides a much more accurate measure of whether or not the job is one that pays for performance. The longitudinal nature of the data also enables us to control for worker-specific fixed effects and show that performance pay is not merely a "label" for being a highly productive worker.

Our empirical results confirm that wages are more closely linked to both observed (education, etc.) and unobserved (worker-specific fixed effects) worker characteristics in performance-pay than in non-performance-pay jobs. We then illustrate the importance of performance pay in the growth in wage inequality by contrasting the actual distribution of wages to the counterfactual distribution that would have prevailed in the absence of performance pay. We also show that wage dispersion has risen faster in performance-pay jobs than in other jobs over this period. This particular finding supports the view that the underlying distribution of individual productivity has become more unequal over time because of changes in the relative demand for different types of workers possibly due to technological change.

Putting together those observations with the fact that the incidence of performance pay has increased over the same time period, we find that, absent performance pay, the variance of log wages would have grown by 21 percent less between the late 1970s and early 1990s. More interestingly, we also find that almost all of the difference between actual wage changes and those predicted in the absence of performance pay occurs at the top end of the wage distribution. In particular, we find that much less of the dramatic growth in wage inequality above the 80th percentile would have occurred in the absence of performance pay.

However, for the reasons discussed above, it would be premature to claim that the growth in performance pay explains 21 percent of the growth in the variance of wages, and most of the increase in inequality above the 80th percentile. We can, nonetheless, infer that performance pay is, at a minimum, a very important channel through which other underlying sources of changes in the distribution of worker productivity, such as skill-biased technical change (SBTC), have been translated into higher wage inequality, especially at the top end of the distribution. Absent this channel, inequality would have increased substantially less between the late 1970s and the early 1990s.

The paper proceeds as follow. In Section II, we present some background on performance-pay schemes, and propose a simple model built upon the insight of Lazear (1986) that the reason performance pay is used is because at the time a worker is employed one cannot observe her ability. Two important predictions of this model are that more productive workers are those who tend to be paid for performance, and that an increase in the return to ability results in more firms choosing to use performance pay. In Section III, we present our empirical model and the testable implications of the theoretical model. In Section IV, we present the data used for the empirical analysis and illustrate the growth in the incidence of performance pay over time. Section V presents the main estimates of the effect of performance pay on the wage structure. We then show in Section VI the connection between performance pay and the growth in wage inequality between the late 1970s and the early 1990s, and conclude our discussion in Section VII.

II. Performance Pay

In the standard competitive model firms and the rest of the labor market observe the marginal product of workers, while competition ensures that the wage is equal to a worker's marginal product. In this setting, modes of payment (fixed wages, performance pay, etc.) have no empirical content since no matter how workers are paid, they are paid for their marginal product. In practice, firms appear to find the problem of setting wages equal to marginal products difficult if not daunting.[2] Over the last thirty years, the economics literature has

[2]Stephen Kerr (1975), in a paper that has earned a place in the canonical MBA course on human resource management, provides a number of examples of firms that, in his opinion, completely fail in their attempt to encourage and pay people according to their marginal product. See also Gibbons (1997), page 9.

explored a number of reasons why firms may not be able to implement pay-for-performance systems, most of which center around monitoring costs. In this section, we provide a brief overview of key empirical findings about the determinants of the incidence of performance pay in the labor market. We then present a simple model that provides the main conceptual framework underlying our empirical analysis.

II.A. Why do Firms Use Performance Pay?

There are a number of reasons why it may be in the interest of firms to introduce performance-pay schemes, even if this entails substantial monitoring and administrative costs. As always, firms will be willing to incur these additional costs provided that they obtain sufficient benefits in return. A commonly mentioned benefit of performance pay is that it provides incentives for workers to exert more effort. But even if performance pay has no effect on workers' effort, when workers are heterogeneous in terms of their innate productive abilities it can be profitable for firms to pay the monitoring cost and then attract more able workers by paying them a wage that better reflects their productivity. In such a setting, performance pay plays an important role in sorting workers across different jobs and/or employers.[3]

Since the cost of obtaining a good measure of the performance of workers is likely to be related to job characteristics, the incidence of performance-pay schemes should also vary according to these characteristics. This prediction holds regardless of whether performance pay is used for incentive or sorting reasons. Using data from the BLS industry wage survey, Brown (1990) explores how the choice between a fixed salary, merit pay and piece-rate compensation depends on monitoring costs. He finds that firms choose standard rates when monitoring costs are high, as is the case with complex jobs. Merit pay systems are more likely to be used when workers feel that their evaluations are fair.

MacLeod and Parent (1999) consider a similar question using a number of panel data sets to control for unobserved worker-specific characteristics. They also extend Brown's analysis to a broader class of compensation systems, and differentiate between bonus pay, commission contracts, and piece-rate contracts. They find that commission contracts are widely used in sales jobs, where the level of sales provides a clean measure of performance. When performance measures are more subjective, then firms either use bonus pay or pay as a function of hours or days worked, with little explicit pay-for-performance.

[3]See Lazear (2000) for some evidence on worker sorting.

In addition to monitoring costs, there are a number of reasons why performance pay may be chosen over other methods of payment in different jobs. Firms that employ high-turnover workers may be more likely to introduce performance-pay schemes than firms with a more stable workforce that can rely upon deferred payments (promotions, pension plans, etc.) to tailor compensation to the characteristics of workers. Indeed, Goldin (1986) shows that around the turn of the 20th century, piece-rates were more widely used in female- than male-dominated occupations, a phenomenon she attributes to the fact that female workers had a higher rate of turnover. Interestingly, piece-rates were more widely used back then than they are today. As modern management practices were introduced and the fraction of clerical and managerial workers grew steadily over time, long-term employment relationships became more prevalent and firms started relying on promotions and other schemes instead of performance pay to provide incentives to their workers.

II.B. Performance Pay and Wage Inequality

As the above discussion makes clear, the decision of firms to introduce performance pay potentially depends on a large number of factors. It is also clear that a decrease in monitoring (or related information processing) costs always increases the probability that firms will use performance pay instead of fixed wages, regardless of the precise reason why firms use performance pay. One would also expect performance pay to increase wage dispersion relative to a payment system based on fixed wages. This can be trivially seen in the case of a firm that pays all workers the same fixed wage when it does not have any information on the ability or the actual performance of individual workers, while differences in productivity are rewarded in a firm that uses performance pay.

Based on these two predictions, it is tempting to propose a simple explanation for how performance pay has contributed to the recent increase in wage inequality. As is well known, the cost of collecting, processing, and analyzing information has declined over time with advances in information and communication technologies. As a result, the cost of introducing performance-pay schemes that require collecting and processing information about workers' performance has presumably declined too, resulting in a growth in the incidence of performance pay. Combining this with the idea that performance pay increases wage inequality, it follows that the growth in performance-pay jobs should have contributed to the rise in wage inequality in the United States.

However, there are a number of reasons why this story is overly simplistic. First, even if performance pay increases wage dispersion among workers who are being paid for performance, the overall impact of performance pay also depends on where these workers are in the skill distribution. This is reminiscent of the case of unions and wage inequality, where unions may end up increasing overall inequality by creating a wedge between union and non-union workers that offsets the equalizing effect of unions within the union sector.[4]

Second, there are good reasons to believe that SBTC, or other explanations that have been suggested for the growth in wage inequality, also have an impact on the decision of firms to use performance pay. In particular, in the sorting model of performance pay discussed above, as the productivity gap between more- and less-skilled workers increases, it becomes more and more advantageous for firms to introduce performance pay to distinguish highly productive workers from less productive workers.

Third, changes in the market for top executives, where performance pay has always been widespread and has also been growing over time (Frydman and Saks (2007)), are hard to reconcile with a simple story based on declining monitoring costs, or in related costs of designing sophisticated compensation systems based on stock options, etc. In contrast, the growth in the share of stock options in total compensation is consistent with the market model of Gabaix and Landier (2008) where performance pay is used for selection purposes. Note that these changes are also consistent with "skimming" stories where executives use performance pay as a cover for rent extraction (e.g., Bebchuk and Fried (2004) and Bertrand and Mullainathan (2001)).

II.C. Model

We now explore these issues more formally using a model of performance pay presented in detail in Appendix 1 in the supplemental material to the paper (we refer to the supplemental material (Lemieux, MacLeod and Parent (2008)) as the Web Appendix from hereinafter). The model builds upon Lazear (1986)'s observation that the reason performance pay is used is because at the time a worker is hired the employer cannot observe her ability. This may result in a mis-match between what the worker is capable of doing and what the employer expects. Linking compensation to performance can reduce this mis-match, and thereby

[4]For example, unions reduce wage inequality among men but not among women for whom unionization is concentrated in the upper end of the skill distribution (Card, Lemieux, and Riddell (2004)).

increase overall productivity. However, the introduction of an effective performance pay system is expensive, and thus one faces a trade-off between the cost of introducing such a system, and the benefits in terms of improved match quality.

Suppose a worker i paid a wage w_{ij} for job j obtains utility $U_{ij} = w_{ij} - exp\left(e_{ij} - \alpha_i\right)$, where e_{ij} is effort and where ability is given by the latent variable, $\alpha_i \sim N\left(\hat{\alpha}_i, \sigma_i^2\right)$. What we call "effort" here can be more broadly interpreted as the effective skills supplied by the worker to complete some specific tasks or duties. For example, workers with lower levels of education (lower α) can supply the same effective skill and perform the same tasks as more educated workers, but doing so is more expensive in utility terms. It is assumed that conditional upon worker characteristics x_i, the mean and variance are known and given by: $\hat{\alpha}_i = E\left\{\alpha | x_i\right\}$ and $\sigma_i^2 = var\left\{\alpha | x_i\right\}$. Following a long standing tradition in labor economics (Jovanovic (1979) and Harris and Holmström (1982)) it is assumed that information is symmetric; both the worker and firm learn α_i at the same time.

Output y_{ij} is assumed to be a linear function of effort:

$$y_{ij} = k_j + \beta\gamma_j e_{ij},$$

where k_j is the output produced on job j regardless of effort and γ_j is the marginal product of effort on job j. The parameter β represents a market return to effort linked, for instance, to the degree of skill bias in technology. Under performance-pay contracts, net output is obtained by subtracting the cost of monitoring effort, M_j.

Under fixed-wage contracts, workers agree to supply a fixed level of effort $\bar{e}_{ij}$ in exchange for a wage w_{ij}^{FW}. Under performance-pay contracts, the firm and the worker agree to a contract linking the wage w_{ij}^{PP} to effort, and the worker sets her effort e_{ij} optimally once her ability α_i is revealed. As mentioned above, we can think of effort as the tasks or duties performed by a worker on a job. For fixed-wage jobs, the worker and the firm agree on specific duties to be performed in exchange for a fixed wage. For performance-pay jobs, a worker is free to pick the tasks or duties that maximize utility. Firms simply design a contract to make sure the interests of the worker are aligned with those of the firm. Once this is done, there is no need to specify strict duties to be performed, and productivity is improved by letting workers tailor their duties to their own skills and abilities.

We show in Appendix 1 that under a fixed-wage contract, the wage is:

$$w_{ij}^{FW} = m_j + \beta\gamma_j\left(\hat{\alpha}_i - \sigma_i^2\right), \quad (1)$$

where $m_j = k_j + \beta\gamma_j log(\beta\gamma_j)$. Under a performance-pay contract, the observed wage is given by:

$$w_{ij}^{PP} = m_j + \beta\gamma_j\alpha_i - M_j, \quad (2)$$

while the *ex ante* expected wage, $\hat{w}_{ij}^{PP}$, conditional on observed characteristics x_i, is the same as above except that the actual value of ability, α_i, is replaced by its expected value, $\hat{\alpha}_i$.

Proposition 1 in Appendix 1 shows that in a match between worker i and firm j, a performance-pay contract is used if and only if $\hat{w}_{ij}^{PP} \geq w_{ij}^{FW}$, or whenever the selection rule

$$\beta\gamma_j\sigma_i^2 \geq M_j \quad (3)$$

is satisfied. Thus, performance-pay contracts are chosen whenever the efficiency gain of performance pay, $\beta\gamma_j\sigma_i^2$, exceeds its cost, M_j. The efficiency gain grows with the conditional variance of ability, σ_i^2, because performance-pay jobs more closely tailor workers' abilities to their work efforts. In contrast, mis-match in fixed-wage jobs rises with σ_i^2. This effect is magnified by the extent of the return to effort on the job (γ_j) or in the overall market (β).

This simple selection rule provides a number of interesting predictions about the conditions under which performance pay is chosen over fixed wage contracts. Obviously, reducing the monitoring costs M_j increases the likelihood of selecting performance pay. Jobs, like executive positions, where output is more sensitive to effort (high γ_j) are also more likely to offer performance pay. Similarly, if β increases because of SBTC, so will the likelihood of choosing performance pay over fixed wages. Finally, performance-pay contracts are more likely to be selected for workers with a higher conditional variance of ability, σ_i^2. Since it is well known that the within-group variance of wages grows with education (see, e.g., Lemieux (2006)), it is reasonable to assume that σ_i^2 is a growing function of expected ability, $\hat{\alpha}_i$.

Figures I and II illustrate some basic implications of the model. For the sake of simplicity, we assume that σ_i^2 is a linear function of expected ability, $\hat{\alpha}_i$: $\sigma_i^2 = \delta\hat{\alpha}_i$. Substituting into equation (3), it follows that performance pay is chosen whenever $\hat{\alpha}_i \geq M_j/\delta\beta\gamma_j$. Here we condition on a specific job j, and discuss the case with multiple jobs in the next section. Figures I and II show that performance-pay workers are concentrated at the top end of the ability distribution. As a result, there is also more wage inequality due to higher returns to observed ability among these workers at the top end (performance-pay workers) than among workers at the low end (fixed-wage workers) of the distribution. Figure I then shows what happens when monitoring costs are reduced from M to M'. The fraction of performance-pay workers increases, and so does inequality since wages at the very top end increase, while wages at the bottom end (fixed wage jobs) remain constant.

A very different explanation for the growth in performance pay illustrated in Figure II is that an increase in returns to effort (or skill) from β to β' induces more firms to switch to performance pay. Unlike Figure I, Figure II shows that the return to ability increases for fixed-wage jobs and increases even more for performance-pay jobs. This is an important and testable difference between the two scenarios illustrated in Figures I and II that we will examine in detail in Section V.

The two scenarios illustrated in Figures I and II have very different implications for the nature of the connection between the growth in performance pay and the growth in wage inequality. In Figure I, the growth in performance pay results in an increase in inequality only to the extent that it moves workers from a less unequal (fixed wages) to a more unequal (performance pay) sector. In Figure II performance pay also interacts with the underlying growth in β because a given increase in β has a larger impact on the return to ability in performance-pay than in fixed-wage jobs. In that sense, performance pay provides an additional channel through which underlying changes in the relative productivities of different groups of workers (such as SBTC) get translated into higher inequality at the top end of the distribution.

III. Empirical Model and Testable Implications

The two wage equations (1) and (2) provide a number of interesting testable implications. For the sake of simplicity, we still maintain the above assumption that $\sigma_i^2 = \delta\hat{\alpha}_i$, where $\delta > 0$. The wage equation under fixed wages becomes

$$w_{ij}^{FW} = m_j + \beta\gamma_j(1-\delta)\hat{\alpha}_i, \tag{4}$$

while the wage equation under performance pay can be rewritten as

$$w_{ij}^{PP} = (m_j - M_j) + \beta\gamma_j\hat{\alpha}_i + \beta\gamma_j(\alpha_i - \hat{\alpha}_i). \tag{5}$$

There are three key differences between these two wage equations, conditional on a job j. First, the intercept is lower for performance-pay than fixed-wage contracts because of the fixed monitoring cost, M_j. Second, the return to expected ability $\hat{\alpha}_i$ is larger under performance pay than fixed wages, which explains why high-ability workers sort themselves into performance pay. Third, there is an error component linked to unobserved ability $(\beta\gamma_j(\alpha_i - \hat{\alpha}_i))$ under performance pay, but not under fixed wages.

All these implications are obtained conditional on a job j. In Appendix 1, we also discuss the market equilibrium in the case where workers with observed characteristics x_i have the choice between different jobs j. We show that, in the simplest version of the model, the "job effects" on wages linked to either observed (industry and occupation) or unobserved (employer-employee job match) are similar in performance-pay and non-performance-pay jobs. We also show that these job effects should be less important in performance-pay than in non-performance-pay jobs in a more realistic setting where *i)* workers partly sort themselves into different jobs on the basis of their unobservable ability (as in Gibbons et al. (2005)), and *ii)* search costs prevent firms from exactly tailoring a fixed wage job to the precise characteristics x_i of each worker.

We now summarize these various predictions using general empirical specifications of the wage equations for the two types of jobs. As a matter of notational convention we use the superscript p for performance-pay jobs, and n for "non-performance-pay jobs" (i.e. fixed-wage jobs). The wage equation for worker i on job j at time t under performance pay is

$$w_{ijt}^{p} = a_t^p + x_{it}b_t^p + z_{ijt}c_t^p + d_t^p\theta_i + \nu_{ij}^p + \varepsilon_{ijt}^p,$$

while the wage for non-performance-pay jobs is

$$w_{ijt}^{n} = a_t^n + x_{it}b_t^n + z_{ijt}c_t^n + d_t^n\theta_i + \nu_{ij}^n + \varepsilon_{ijt}^n,$$

where x_{it} represents standard observable worker characteristics like potential experience and education; $\theta_i = \alpha_i - \hat{\alpha}_i$ is the unobservable ability component; z_{ijt} is a set of observed job characteristics like occupation or industry; ν_{ij}^p and ν_{ij}^n are "firm-specific" wage components; and ε_{ijt}^p and ε_{ijt}^n are idiosyncratic error terms.

The main empirical implications discussed above are summarized as follows:

1. The wage intercept is lower in performance-pay than non-performance-pay jobs: $a_t^p < a_t^n$

2. The return to observable worker characteristics, x_{it}, is larger in performance-pay jobs than non-performance-pay jobs: $b_t^p > b_t^n$

3. The return to observable job characteristics, z_{ijt}, is smaller in performance-pay than non-performance-pay jobs: $c_t^p < c_t^n$

4. The return to unobservable ability θ_i is larger in performance-pay jobs than non-performance-pay jobs: $d_t^p > d_t^n$. Although the model predicts that $d_t^n = 0$, the estimated value of d_t^n will be positive if the market observes some part of θ_i (e.g. the quality of education, past productivity, etc.) that is not reflected in observable characteristics x_{it}, as in Gibbons et al. (2005).

5. The variance of the firm-specific component is smaller in performance-pay than non-performance-pay jobs: $var(\nu_{ij}^p) < var(\nu_{ij}^n)$

IV. Data

The bulk of our analysis is conducted using data from the PSID. The main advantage of the PSID is that it provides a representative sample of the workforce for a relatively long time period, essential for studying the effect of performance pay on wage inequality. One disadvantage of the PSID is that our constructed measures of performance pay are relatively crude for reasons discussed below. To probe the robustness of the results based on the PSID, we re-estimate some of the key models using the National Longitudinal Survey of Youth (NLSY).

IV.A. The Panel Study of Income Dynamics (1976-1998)

The PSID sample we use consists of male heads of households aged 18 to 65 with average hourly earnings between \$1.50 and \$100.00 (in 1979 dollars) for the years 1976-1998, where the hourly wage rate is obtained by dividing total labor earnings from all jobs by total hours of work, both reported retrospectively for the previous calendar year.[5,6] Given our focus on performance pay, this wage measure based on total yearly earnings, inclusive of performance pay, is preferable to "point-in-time" wage measures that would likely miss infrequent payments (e.g. bonuses) of performance pay.[7]

Individuals who are self-employed are excluded from the analysis since our measure of performance pay based on receiving bonuses, commissions, or piece-rates is defined for employed workers only.[8] We also exclude workers from the public sector since it is not clear what it means to pay workers for their productivity in a sector where employment and wage setting decisions are not based on profit maximization (we show in Table B.1 of the Web Appendix that including public sector workers has little impact on the results). This leaves us with a total sample of 26,146 observations for 3,053 workers. All of the estimates reported in the paper are weighted using the PSID sample weights.

Identifying Performance Pay In the PSID, we construct a performance-pay indicator variable by looking at whether part of a worker's total compensation includes a variable pay

[5]In the PSID, data on hours worked during year t, as well as on total labor earnings, bonuses/commissions/overtime income, and overtime hours, are asked in interview year t+1. Thus we actually use data covering interview years 1976-1999. Annual earnings were top coded at $99,999 until 1982 (and not top coded since then), but only a handful of individuals were at the top code. We trim very high values of wages (above $100.00 in 1979 dollars) but do not otherwise adjust for top coding.

[6]Our focus on male heads of households stems from the fact that only heads are asked about their income derived from bonuses, commissions, or overtime. In the PSID, males are designated as the head in all husband-wife pairs. The same is true if the female has a boyfriend with whom she has been living for at least a year, even if the female is the person with the most financial responsibility in the family unit. Consequently, the sample of female heads is relatively small. Using the same sample selection criteria as the ones we use for males would leave us with 1,367 females for a total of 8,185 observations. Perhaps more importantly, issues of representativeness would arise as those female heads are disproportionately nonwhite (24.4 percent) and are much less likely to be married (9.2 percent).

[7]See Lemieux (2006) and Autor, Katz, and Kearney (2005) for a detailed comparison of these two types of wage measures using March (average hourly earnings) and MORG (point-in-time wage) CPS data. As in the CPS, we find that inequality is lower in the PSID when a point-in-time wage measure is used instead of the average hourly earnings measure used throughout the paper. For workers not paid for performance, the difference in standard deviations is 0.015 for salaried workers compared to 0.064 for workers paid by the hour. We also find that the difference between the two measures is larger for salaried workers paid for performance (0.036) than for those not paid for performance (0.015). This confirms that inequality using point-in-time wage measures may slightly understate inequality as it misses the contribution of performance pay. But since hourly workers are unlikely to be paid for performance (see Table I), performance pay cannot account for much of the difference in inequality between the two wage measures for this large group of workers. Finally, trends in inequality based on the two measures are generally very similar.

[8]Self-employed workers can be viewed as being, by definition, paid for performance regardless of the mode of payment (earnings, dividends, etc.) they use to remunerate themselves.

component (bonus, commission, or piece-rate). For interview years 1976-1992, we are able to determine whether a worker received a bonus or a commission over the previous calendar year through the use of multiple questions. First, workers are asked the amount of money they received from working overtime, from commissions, or from bonuses paid by the employer.[9] Second, we sometimes know only whether or not workers worked overtime, and if they are working overtime in a given year, not the amount of pay they received for overtime. Thus, we classify workers as not having had a variable pay component if they worked overtime. Third, workers not paid exclusively by the hour, or not exclusively by a salary, are asked how they are paid: they can report being paid commissions, piece-rates, etc., as well as a combination of salaried/hourly pay along with piece-rates or commissions.[10] Through this combination of questions, we are thus able to identify *all* non-overtime workers who received performance pay in bonus, commission, or piece-rate form.

Starting with interview year 1993, there are separate questions about the amounts earned in bonuses, commissions, tips, and overtime for the previous calendar year. Thus, there is no need to back out an estimate of bonuses from an aggregate amount since the question is asked directly. For the sake of comparability with the pre-1993 years, we nevertheless classify as receiving no performance pay all workers who report any overtime work. In this way we are able to determine whether a worker's total compensation included a performance-pay component for each year of the survey. One obvious drawback is that it is likely that the performance-pay component we construct will be noisy for hourly workers, though not for salaried workers who are not eligible for overtime payments. However, due to our treatment of overtime workers, we conservatively lean on the side of misclassifying workers as receiving no performance pay even when they do.[11]

[9] Note that the question refers specifically to any amounts earned from bonuses, overtime, or commissions in addition to wages and salaries earned.

[10] In many survey years workers are not asked if their compensation package involves a mixture of salary/hourly pay and a variable component. All they are asked is how they are paid if not by the hour or with a salary. Although there is no way to directly verify it, this likely results in understating the incidence of any form of variable pay because workers are not allowed to answer that they are paid, say, a salary, and then report a commission: they have to choose. Our assertion that this response likely understates the extent of variable pay is motivated in part by the fact that workers in the NLSY, to be described below, are not restricted in describing the way they are paid. We find that workers in the NLSY are more likely to report having part of their compensation package contain a performance-pay component.

[11] In an earlier version of the paper, we re-did the analysis for 1992 to 1998 using the finer measure of performance pay that allows us to identify the performance-pay status of overtime workers. Doing so had little impact on the results. It only increased the fraction of workers on performance-pay jobs (for 1992-98) by one percentage point, and regression coefficients were essentially unchanged.

Defining Performance-pay Jobs We define performance-pay jobs as employment relationships in which part of the worker's total compensation includes a variable pay component (bonus, a commission, piece-rate) at least once during the course of the relationship.[12] Since we use actual payments of bonuses, commissions or piece rates to identify performance-pay jobs, we are likely to misclassify performance-pay jobs as non-performance-pay jobs if some employment relationships are either terminated before performance pay is received, or partly unobserved for being out of our sample range. This source of measurement error is problematic because of an "end-point" problem in the PSID data. Given our definition of performance-pay jobs, we may mechanically understate the fraction of workers in such jobs at the beginning of our sample period because most employment relationships observed in 1976 started before 1976, and we do not observe whether or not performance pay was received prior to 1976. Similarly, jobs that started toward the end of the sample period may be performance-pay jobs but are classified otherwise because they have not lasted long enough for performance pay to be observed.

The problem is that, conditional on job duration, we tend to observe a given job match fewer times at the two ends of our sample period than in the middle of the sample. Consider, for example, the case of a job that lasts for five years. For jobs that last from 1985 to 1989, all five observations on this job match are captured in our PSID sample. For jobs that last from 1973 to 1977, however, only two of the five years of the job match are observed, which mechanically reduces the probability of classifying the job as one with performance pay.

Because of this end-point problem, we get an unbalanced distribution of the number of times job matches are observed at different points of the sample period. One simple solution to the problem is to "rebalance" the sample using regression or other methods. In practice, we adjust measures of the incidence of performance pay over time by estimating a linear probability model in which dummies for calendar years and for the number of times the job-match is observed are included as regressors (estimating a logit gave almost identical results). We then compute an adjusted measure of the incidence of performance pay by holding the distribution of the number of times the job-match is observed to its average value for the years 1982 to 1990, which are relatively unaffected by the end-point problem.

[12] We use "jobs", "employment relationship", and "job match" interchangeably. Although the PSID does have information on tenure in the position in most of the survey years spanning the sample period, we do not use it. As is well known, simply determining employer tenure in the PSID can be problematic (see Brown and Light (1992)). As a result, what we call a "job match" could be called an "employer match" instead. We generally use the word "job" for the sake of simplicity.

The end-point problem could also affect the estimates of the effect of performance pay on both the level and the dispersion of wages because the sample of non-performance-pay jobs is being contaminated by observations from performance-pay jobs for which performance-based payments are never observed. We have investigated this issue in detail using a parametric measurement model described in Appendix 2 of the Web Appendix and concluded that, if anything, this measurement problem biases downward the estimated effect of performance pay on the wage structure. For the sake of clarity and simplicity, the wage results we report in the next sections are unadjusted for these measurement issues.

IV.B. Descriptive Statistics from the PSID

Table I compares the mean characteristics of workers on performance-pay and non-performance-pay jobs, respectively. First, notice that 37 percent of the 26,146 observations are in performance-pay jobs.[13] Workers on performance-pay jobs tend to earn more and have higher levels of education than workers on non-performance-pay jobs. Note that the hourly wage rate includes both regular wage and salary earnings and performance pay in the case of workers on performance-pay jobs. Annual hours worked and employer tenure also tend to be higher for workers on performance-pay than non-performance-pay jobs.

The unionization rate (percent covered by a collective bargaining agreements) is much lower among performance-pay workers. This suggests that, as expected, the pay structure in union firms corresponds more closely to the fixed-wage contracts discussed in Section II. Another important difference is that there is a much higher fraction of workers paid by the hour in non-performance-pay than performance-pay jobs. Conversely, workers on performance-pay jobs are more likely to be salaried workers than those on non-performance-pay jobs. This is an important point since the growth in wage inequality has been stronger among salaried than hourly workers (Lemieux (2006)). Performance pay is thus more likely to affect the very group of workers who have experienced the largest increase in inequality, and who are also least likely to be affected by other institutional factors such as the minimum wage or unionization. With the exception of potential experience, the mean characteristics in performance-pay jobs are statistically different from those in non-performance-pay jobs.

An important point illustrated at the bottom of the table is that, of the 3053 workers,

[13] The 37 percent figure is unadjusted. This fraction jumps to 42 percent when we adjust for the end-point problem using the procedure discussed above (see the lower right corner of Appendix Table A.1).

1271 are observed on a performance-pay job, and 2616 are observed on a non-performance-pay job. So 834 workers (1271+2616-3053) are "switchers" observed on both types of jobs, which is essential for identifying models with fixed effects presented in Section V.

The cross tabulations shown in Table A.1 confirm that performance pay is more prevalent in high-wage occupations like professional, managerial, and sales positions than in other occupations. For example, the fraction of workers on performance-pay jobs ranges from only 30 percent for craftsmen to 78 percent for sales workers. Across industries, the incidence of performance pay ranges from a low of 33 percent in mining and durables to a high of 65 percent in finance, insurance and real estate (FIRE). Note that the (1-digit) industry and occupation categories shown in the table are the ones we use to control for industry and occupation effects in the regression models presented later in the paper.

Figure III presents kernel density estimates of the distribution of wages for performance-pay and non-performance-pay jobs. The figure shows that hourly wages have a higher mean and median, and are less evenly distributed among performance-pay than non-performance-pay jobs.

We next turn to the time trends in the prevalence of performance pay. Figures IVa and IVb show the evolution of the fraction of performance-pay jobs for various subgroups of the workforce. In all cases, we correct for the end-point problem using the procedure described above. Figure IVa shows that the overall incidence of performance-pay jobs has increased from about 35 percent in the late 1970s to around 45 percent in the 1990s. The figure also shows the simpler measure based on the fraction of workers actually reporting performance pay in a given year. This alternative measure clearly understates the incidence of performance-pay jobs since workers on performance-pay jobs will not necessarily receive a performance payment (like a bonus) in each year on the job. One advantage of this simple measure, however, is that it is not affected by the end-point problem and provides additional evidence of the robustness of the underlying trends in performance pay. Indeed, even this crude measure of performance pay clearly increases over time, especially in the 1980s.

Figure IVa also shows the fraction of workers covered by a collective bargaining agreement. Interestingly, the decline in unionization and the growth in performance pay are both concentrated in the same period (the 1980s). This suggests that one way de-unionization may have contributed to the growth in wage inequality is by allowing firms to offer more variable pay.[14] Figure IVb shows, however, that the growth of performance pay is not simply a spurious consequence of the decline in unionization. In particular, the figure shows that the incidence of performance pay has been growing among both union and, especially, nonunion workers.

Figure IVb also reports another way of looking at the increase in the incidence of performance-pay jobs by breaking it down by how workers are paid. The figure shows that the bulk of the increase in performance pay is driven by salaried workers who are, incidentally, less likely to be unionized. By contrast, performance pay is less prevalent and grows more slowly over time among workers paid by the hour. The increase in the incidence of performance-pay jobs among salaried workers is quite remarkable. It increases from less than 45 percent in the late 1970s, to nearly 60 percent by the end of the sample period.

Note that performance pay represents a relatively modest share of total earnings (Figure B.1 in the Web Appendix shows that the median share is 4.4 percent). However, this does not mean that performance pay has a limited impact on total compensation since we expect (and find in Table B.1 of the Web Appendix) the straight wage component to be more sensitive to workers' characteristics on performance-pay than non-performance-pay jobs. In order to pay for performance, the employer must evaluate the worker, which then affects the straight wage through promotions and job assignment. Hence, even though performance pay is a relatively small fraction of compensation for most workers, the fact that it exists is a signal of more careful monitoring.

IV.C. The Growth in Performance Pay: Some Additional Evidence

Two important findings reported in Figure IV are that different measures of performance pay indicate a clear growth in performance pay, and that this growth is not just a spurious consequence of de-unionization. Table II takes a more general look at these issues by considering a number of additional measures of performance-pay jobs, and possible explanations for the growth in performance pay beyond de-unionization.

In this table and the remainder of the paper, we focus on changes between the late 1970s (1976-79) and early 1990s (1990-93). We use 1990-93 as our end period (instead of data up to 1998) to minimize the end-point problems mentioned above, though using 1994-98 yields

[14] See Freeman (1993), Card (1996), and DiNardo, Fortin, and Lemieux (1996) for evidence that de-unionization accounts for about a quarter of the growth in male wage inequality during the 1980s.

similar results. While it would arguably be better to use a base period further away from the first observation year (1976) to reduce end-point problems, by doing so we would miss part of the large increase in inequality (and performance pay) that took place in the early 1980s.

We argued above that it was too restrictive to just classify a worker-year observation as one where the worker is paid for performance when an actual payment (bonus, commission, or piece-rate) is received in the current year. Instead, our preferred measure is whether or not a worker on a given job receives performance pay at any time during the observed employment relationship. One could argue that this alternative definition is "too loose". For example, if we have twenty observations on a worker in a given job, but performance pay is only observed once, it is not clear to what extent such a job is really one that pays for performance. A reasonable alternative is to classify as performance-pay jobs only those for which the frequency of actual performance-based payments exceeds a certain threshold. With this in mind, Table II shows both the incidence and the growth in performance pay under increasingly strict definitions.

Column 1 of Table II shows the results for our preferred measure of performance pay based on payments of bonuses, commissions, or piece-rates in any year of the employment relationship.[15] In column 2, we only classify jobs as performance-pay when a payment is observed at least one time out of five. We increase the minimum intensity to one time out of two in column 3, and then present the simple measure based on actual payment in the current year in column 4.

The most important pattern that emerges from the table is that, regardless of the measure being used, there is always a substantial increase in performance pay between the late 1970s and the early 1990s. In fact, while the incidence of performance pay obviously decreases when stricter measures are considered, the growth is, if anything, larger in relative terms for these stricter measures. We also show in column 5 that essentially all the growth in performance-pay jobs is driven by the bonus-pay component, as opposed to commissions or piece-rates.

The table also shows the impact of the adjustment for the number of times the job-match

[15] Note that among observations defined as performance pay that way, we observe an actual performance payment in 37 percent of cases. The average intensity increases to 57 percent, however, when we average the frequency of payments across jobs, i.e. put an equal weight on all jobs irrespective of the number of observations we have for each job. So even under our broadest measure of performance pay, actual performance payments are frequently observed.

is observed. In the case of the broadest measure reported in column 1, the adjustment reduces the growth in performance pay from 12.9 (row 2) to 7.1 (row 3) percentage points. The reason the adjustment is quite large is that the base period we chose, 1976-79, is more directly affected by the end-point problem than the end period of 1990-93.

The second part of the table shows the contribution of other factors to the growth in performance pay.[16] Using these estimates, we perform a simple decomposition to see by how much the incidence of performance pay would have changed if the different explanatory factors had remained constant over time.

In the case of our main measure of performance pay (column 1), row 5 of Table II shows that about a third (2.5 percentage points) of the 7.1 percentage points increase in performance pay can be linked to changes in these explanatory factors. The most important factor is de-unionization that accounts for 1.4 percentage point of the growth in performance pay, followed by changes in the distribution of industry and occupation that each explain a little more than half of a percentage point. The remaining factors (education, etc.) account for essentially none of the growth in performance pay. The results for other measures of performance pay reported in columns 2 to 5 are very similar to those for our broader measure of performance pay. We conclude from Table II that the growth in performance pay measured in the PSID is very robust to the way performance-pay jobs are defined, and cannot be explained by other factors such as de-unionization.

One additional source of evidence is the NLSY, which asks more explicitly about performance pay in several years starting in 1988.[17] Using a similar sample to the one used for the PSID, we find that the incidence of performance-pay jobs increases from 26.1 percent in the late 1980s to 30 percent in the late 1990s, broadly consistent with the evidence from the PSID. We also looked at another source of information based on a survey of Fortune 1000 corporations conducted between 1987 and 2003 (see Lawler III (2003)). The survey asks firms about the fraction of their workers with some forms of performance pay and reports results in categories such as 0 to 9 percent, 10 to 19 percent, etc. We compute the implied fraction of workers with performance pay using the mid-points of these intervals. The implied fractions are 20.7 in 1987, 27.1 in 1990, 34.7 in 1996, and 44.5 in 2002. Once again,

[16] These contributions are computed by first estimating a linear probability model with a full set of dummies for time periods (1976-79, 1980-84, 1985-89, 1990-93, and 1994-98) and the number of times a job match is observed (1 to 22), as well as dummies for industry (10), occupation (8), marital status, race, union status, a cubic function in potential experience, and a quadratic function in job tenure.

[17] More details on the NLSY data are provided in Appendix 3 of the Web Appendix.

these trends confirm the growth in performance pay measured (imperfectly) in the PSID.

V. The Wage Structure in Performance-pay and Non-performance-pay Jobs

The model of Section II provides a number of testable implications about differences in the structure of wages between performance-pay and non-performance-pay jobs. We now present the estimation results and show that they are consistent with the predictions of the model outlined in Section III.

V.A. Simple Regression Analysis

Table III reports simple regression estimates of the effect of performance pay on wages (full compensation, including the actual performance-based payment). These regressions are provided as a benchmark before moving to the core predictions of the model about the differences in the returns to measured and unmeasured characteristics in the two pay regimes. Since we have repeated observations for the same individual observed in a given job match (the level at which performance-pay jobs are measured), we allow for correlation in the error terms by clustering standard errors at the job-match level in Table III and subsequent tables.

The first column of Table III reports the results of an OLS regression of the log hourly wage on a dummy for performance-pay jobs. The regressions reported in Table III also control for standard worker characteristics x_{it} (years of education, a cubic in potential experience, dummies for race and marital status, and the local unemployment rate) and job characteristics z_{ijt} (union status, a quadratic in seniority, and industry and occupation dummies), though the estimated coefficients for these variables are not reported in the table.

The estimated effect of the performance-pay job dummy is positive (0.087) and statistically significant, though it is much smaller than the raw wage gap reported in Table I (the unadjusted difference in mean log wages is 0.224). The second column shows that the effect of having a performance-pay job declines but remains very significant when a dummy for performance pay received during the year is included. When worker-specific fixed effects are introduced in column 3, the effects of performance-pay jobs and of receiving performance pay in a given year become smaller but remain positive and significant. For both this table and for the other results reported in the paper, the fixed effect models are precisely estimated due to the large number of workers who switch between the two types of jobs (see Table I).

The results are consistent with the positive sorting into performance pay predicted by the model of Section II. Note that introducing observed covariates reduces the wage gap by 0.139 (raw gap of 0.224 compared to OLS estimate of 0.087), compared to a further 0.047 reduction (column 3 vs. column 1) when worker-specific fixed effects are added to the wage equation. This implies that most of the sorting happens on observable dimensions of skills.

Also note that the estimated effect of receiving a performance-based payment in a given year is around 4-4.5 percent in columns 4 and 5 where we further control for worker-job fixed effects (the effects of performance-pay jobs are no longer identified in this specification). This suggests that performance-pay is not merely *displacing* base pay, but results in *increased* compensation, even after controlling for individual and job-specific characteristics.

V.B. Return to Skill in Performance-pay and Non-performance-pay Jobs

Table IV provides a direct test of some of the implications of the model. Columns 1 and 2 report separate estimates of a standard wage equation for performance-pay and non-performance-pay jobs, respectively.[18] The estimated models include the same variables as those included in Table III. We only report, however, the estimated effect of years of education, potential experience and job tenure.[19]

As expected, both the return to education and the return to experience are substantially larger in performance-pay than non-performance-pay jobs (e.g., effect of 0.093 vs. 0.067 for education). We next show in column 3 a more parsimonious specification where we estimate a pooled regression model where education, experience, and tenure (including the full cubic in experience and the quadratic in tenure) are interacted with the performance-pay job dummy,

[18]A more sophisticated approach would be to use the technique of Gibbons et al. (2005) where the return to unobserved ability is allowed to differ across job types (as our model predicts), and learning induces endogenous mobility across jobs. For the sake of simplicity, however, we only control for a standard fixed effect since the results of Gibbons et al. (2005) suggest that, at least for occupations, doing so corrects for most of the endogeneity bias due to the fact that job choice depends on unobserved ability.

[19]To further simplify the table, we only report the effect of 20 years of potential experience and 10 years of job tenure. This is obtained by computing the predicted effect from the polynomial specifications (cubic in experience, quadratic in tenure) at 20 (10) years of experience (tenure). We only report these results since qualitatively similar results were obtained using either 5, 10, and 20 years, and the mean values of experience and tenure are close to 20 and 10 years, respectively (Table I).

while other variables are constrained to have the same effect for both performance-pay and non-performance-pay jobs. Since doing so yields similar results, we keep this specification for the rest of the table so that all of the differences between performance-pay and non-performance-pay jobs are summarized by the interaction terms reported in the table. The pooled models also provide a simple way of testing whether the returns to characteristics are different for performance-pay and non-performance-pay jobs.

We first report OLS estimates of the pooled model as a benchmark in column 3, and then add worker-specific fixed effects in column 4. In both cases, we find that the return to education is significantly larger in performance-pay than non-performance-pay jobs.[20] As predicted in Section III, the intercept is also lower in performance-pay jobs when the interactions are included in the specifications, as in columns 3 and 4. Note, however, that the effect of experience is not significantly different for the two types of jobs when fixed effects are included (column 4).[21]

Unlike education and experience, it is not clear a priori whether job tenure is a pure job characteristic linked to administrative pay levels, or is in part a worker characteristic linked to specific human capital accumulation. Table IV shows that the effect of job tenure is lower in performance-pay than non-performance-pay jobs. This supports the view of tenure as a job characteristic. The difference is no longer significant in the pooled regressions with fixed effects, as reported in column 4.

Another obvious job characteristic to look at is occupational affiliation. Since Gibbons et al. (2005) have shown that including worker-specific fixed effects dramatically reduces the magnitude of the occupation effects, we estimate occupational wage differentials for performance-pay and non-performance-pays jobs by interacting the performance-pay job dummy with occupation dummies in the fixed effect model reported in column 4 of Table IV. Consistent with the predictions of Section III, the standard deviation of the occupation effects is smaller in performance-pay jobs (0.042) than in non-performance-pay jobs (0.044).

[20] It is difficult to interpret the main effect of education in the model with worker-specific fixed effects because education is almost time-invariant (for a given person) in our PSID sample. This means that it is difficult to separately identify the effect of education from the fixed effect when running separate models for performance-pay and non-performance pay jobs. The interaction term between performance-pay and education is still identified, however, because of the "switchers" who are observed in both performance-pay and non-performance-pay jobs.

[21] While the interaction term between the performance-pay dummy and experience is not significant at the specific level of experience we look at (20 years), a joint test indicates that the whole experience profile (linear, quadratic, and cubic terms) is significantly different for the two types of jobs in columns 3 and 5.

The last two columns of Table IV allow the return to education to vary over time, in both performance-pay and non-performance-pay jobs.[22] We divide the sample into five periods (1976-79, 1980-84, 1985-89, 1990-93, and 1994-98) and interact the period dummies with years of education, performance pay, and the interaction of these two variables. Given that the growth in the return to education (and wage inequality, more generally) is concentrated in the 1980s, we only report the results for the period 1990-93 in the table. Note that the main effect of education (and of the interaction between education and performance pay) now corresponds to the base period (1976-79). The OLS estimates in column 5 show that, as expected, the return to education increased between 1976-79 and 1990-93, and increased even faster for performance-pay jobs. The coefficient estimates indicate that the return to education increased by 0.0161 for non-performance-pay jobs, and by 0.0351 for performance pay jobs (0.0161 plus 0.0190). The changes are even more pronounced and highly significant when fixed effects are included in column 6. The fact that the returns to skill are increasing faster in performance-pay than non-performance-pay jobs is consistent with the case illustrated in Figure II, where an increase in the relative demand for skilled labor may also be the reason for the growth in performance pay.

V.C. Variance Components Analysis

Having established that observable worker characteristics matter relatively more for performance- than non-performance-pay jobs, while the reverse is true for observable job characteristics, we now look at whether this pattern of results also holds in the case of unobservable characteristics. We do so by performing a variance components analysis on the residuals from the wage regressions estimated separately for performance-pay and non-performance-pay jobs (columns 1 and 2 of Table IV). Going back to the wage equations of Section III, the residual for performance-pay jobs, e^p_{ijt}, is

$$e^p_{ijt} = d^p_t \theta_i + \nu^p_{ij} + \varepsilon^p_{ijt}, \tag{6}$$

while the residual for non-performance-pay jobs, e^n_{ijt}, is

$$e^n_{ijt} = d^n_t \theta_i + \nu^n_{ij} + \varepsilon^n_{ijt}. \tag{7}$$

[22] We also looked at the changes in the returns to other characteristics over time, but education was the only variable for which we systematically found a growing effect.

The parameters of interest to be estimated are the variances of each of the six error components in equations (6) and (7). We estimate the model under the simplifying assumption that the idiosyncratic error terms ε_{ijt}^{p} and ε_{ijt}^{n} are uncorrelated over time. Following Parent (2002), we estimate the variance components by fitting regression models to all the cross-products of residuals for the same individual.[23] This procedure is similar to the equally-weighted minimum distance approach of Abowd and Card (1989), but provides an easy way of dealing with an unbalanced sample like ours.

We first report in panel A of Table V the results estimated over the whole sample. One potential pitfall of using the whole sample is that some individuals are only observed on performance-pay jobs, while others are only observed on non-performance-pay jobs. As a result, the variance of the worker-specific effect θ_i may not be the same in the two subsamples, and differences between the estimated variance components $var(d_t^p\theta_i)$ and $var(d_t^n\theta_i)$ may reflect composition effects related to θ_i, as opposed to true differences in the return to unobservables d_t^n and d_t^p. To control for this potential problem, we report in Panel B the results for the subsample of "switchers" who are observed on both performance-pay and non-performance-pay jobs.

As a benchmark, we start with simple models in columns 1 and 4 where we do not include the variance component linked to the job-match, and also constrain the variance components to be constant over time. We then add the job-match component in columns 2 and 5. In columns 3 and 6, we let the variance of the idiosyncratic terms ε_{ijt}^{p} and ε_{ijt}^{n} and the return to the worker component (the factor loadings) d_t^p and d_t^n change over the five subperiods used in Table IV. The results in the two panels of Table V are very similar, but we focus the discussion on Panel B for the reasons mentioned above. Note that since a large number of cross-products are available (between 32,476 and 99,554 in the different models reported in Table V), the parameters are precisely estimated and are, unless otherwise indicated, statistically different for performance-pay and non-performance-pay jobs.

The results show that, as expected, the worker-specific component θ_i accounts for more of the variation of wages in performance-pay than non-performance-pay jobs. When the job-match term is included in column 2 and 5, the estimated variances of the worker component are 0.102 and 0.053, respectively. Since the ratio of $var(d^p\theta_i)$ and $var(d^n\theta_i)$ is equal to the

[23]See Parent (1999) for a related analysis with the NLSY comparing piece-rate/commission workers and those receiving bonuses to salaried and hourly paid workers. More details on the identification and estimation of the variance components models are provided in the Web Appendix (Appendix 4).

square of the relative return in performance- and non-performance pay jobs, this implies that d^p/d^n is equal to 1.39. In other words, d^p is 39 percent larger than d^n.

Also consistent with predictions, the results indicate that the variance of the job-match component is much smaller in performance-pay (0.002) than non-performance-pay jobs (0.026). In intuitive terms, this suggests that the firm an individual works for explains quite a bit of the wage variation in non-performance-pay jobs, but much less in performance-pay jobs.[24] The variance of the idiosyncratic term that represents year-to-year volatility in wages is slightly smaller in performance-pay than in non-performance-pay jobs, though the model does not have specific predictions in this regard.

In columns 3 and 6 where the variance components are allowed to change over time periods, the factor loadings d_t^p and d_t^n grow for both performance-pay and non-performance-pay jobs. The variance of the worker component shown on the first row now refers to the variance in the base period (1976-79). Consistent with Baker (1997), the factor loading on the person-specific component increases over time. For both performance-pay and non-performance-pay jobs, the factor loadings in Panels A and B are 17 to 31 percent higher in 1990-93 then in 1976-79.

Although the *relative* growth in the factor loadings d_t^p and d_t^n is not statistically different for the two types of jobs, the resulting growth in the variance associated with the worker-specific component is larger for performance-pay jobs because the variance (and the corresponding factor loading) is larger in the base period. For performance-pay jobs, the variance grows from 0.067 in 1976-79 to 0.115 (0.067 times the square of 1.312, the factor loading) in 1990-93, a 0.049 increase. For non-performance-pay jobs, the variance grows from 0.036 to 0.062, a 0.026 increase.

V.D. Robustness Checks

As discussed in Section IV, while our measure of performance pay is rather crude, the growth in performance pay is robust to the way we measure it (Table II). The results for

[24] As pointed out in Section III, due to the greater complementarity between the job effect and worker ability in performance-pay jobs, we would expect that a substantial part of the job-match effect would be absorbed by the individual ability term α_i. Although the results are not shown to save space, this is exactly what happens. When we first fit the covariance structure models with only a job-match term and an idiosyncratic term, the estimates of the job-match terms are very similar in both types of jobs. However, as we can see in Table V, controlling for the worker-specific fixed effect results in the job-match effect becoming negligible in performance-pay jobs.

the wage equations described above are also highly robust to these measurement issues, and to a number of other specification choices. We present a detailed analysis of these robustness issues in the Web Appendix (Table B.1), and only summarize the main findings here. The focus of the robustness analysis is the difference in the returns to education between performance-pay and non-performance-pay jobs, which is the simplest and clearest way of showing the key difference in the wage structure between the two types of jobs.

The first set of alternative specifications reported in the Web Appendix is based on alternative measures of performance pay like the ones reported in Table II. We also look at what happens when public sector workers are included, when richer sets of interactions are introduced, and when only the base wage (net of performance-based payments) is used as the dependent variable. A final estimator is based on a measurement error correction that accounts for the fact that we are more likely to misclassify performance-pay jobs as non-performance pay jobs when we only have a few observations on a given job-match. Using these alternative specifications has little impact on the results. For instance, the average OLS estimate for the eleven additional specifications in Table B.1 is 0.0388, compared to 0.0365 in Table IV (column 3). The average fixed effect estimate is also very similar (0.0141) to the estimate reported in Table IV (0.0165 in column 4).

A second piece of evidence in support of our main findings comes from the NLSY data. As in the case of the PSID, we run separate wage regressions for performance-pay and non-performance-pay jobs. We also exploit the fact that the Armed Forces Qualifying Test (AFQT) score, which is available in the NLSY, can be used as a proxy for unobserved productive characteristics. The results, reported in Table B.2 of the Web Appendix, show that returns to productive worker characteristics (education, experience, and the AFQT score) are larger in performance-pay than non-performance-pay jobs.

VI. Performance Pay and Increasing Wage Inequality

We now return to the main question addressed in this paper: what is the relationship between the growth in performance pay and wage inequality? We begin by presenting the results of simple decomposition, or accounting, exercises. We then discuss the interpretation of these results in light of the possible explanations of the growth in performance pay presented in Section II.

VI.A. Reweighting Estimates

Quantifying the contribution of the change in a wage-determining factor such as performance pay on the wage distribution is a well-known problem in the inequality literature. For example, DiNardo and Lemieux (1997) contrast the observed change in the distribution of wages to the change that would have prevailed in the absence of unions. To do so, they use the reweighting approach of DiNardo, Fortin, and Lemieux (1996) to control for differences in the distribution of worker characteristics in the union and non-union sectors. Just like a standard regression provides a way of adjusting differences in mean wages between two groups for differences in worker characteristics, the reweighting procedure allows to do so for any feature of the wage distribution, and not just the mean.

To fix ideas, let PPJ be a dummy variable indicating whether a worker holds a performance-pay ($PPJ = 1$) or a non-performance-pay ($PPJ = 0$) job. Let X now represent all observable characteristics (both the worker and job characteristics discussed earlier). Following DiNardo and Lemieux (1997), the counterfactual distribution of wages that would prevail if all workers were paid like workers in non-performance-pay jobs can be estimated by reweighting all non-performance-pay observations using the reweighting factor $\psi(X) = \Pr(PPJ = 0)/\Pr(PPJ = 0|X)$. The idea is very simple. Groups like sales workers that are very likely to be paid for performance ($\Pr(PPJ = 0|X)$ is low) will be underrepresented among non-performance-pay workers. So this group has to be given a larger weight to get a distribution of non-performance pay workers that is representative of the whole workforce. This is achieved using the reweighting factor $\psi(X)$, which is large for this group since its denominator, $\Pr(PPJ = 0|X)$, is low. It is easy to estimate the conditional probability $\Pr(PPJ = 0|X)$ by running a simple probit or logit model for the probability of being paid for performance as a function of the observable characteristics X.[25]

Before presenting the decomposition results, we first report some descriptive information on the trends in wage inequality to be explained. Figure V summarizes the changes in wage inequality in our PSID data by showing the evolution of the standard deviation of wages (3-year moving average) in performance-pay, non-performance-pay, and all jobs between 1977

[25] We use a probit model with a more flexible specification than the one used in the wage equations reported in Tables III and IV. Relative to these specifications, we add a set of four education dummies that we also interact with potential experience (linear term), union status, and the race dummy. We also add a cubic in tenure, a dummy for full-time/full-year workers, and an interaction between potential experience and the race dummy.

and 1996. In Figure V and the rest of this section, we follow DiNardo, Fortin, and Lemieux (1996) and weight observations by the numbers of hours worked during the year to get a representative distribution of wages paid over all hours worked in the labor market. As before, we also weight observations using the PSID sample weight.

As expected, the figure indicates a substantial increase in inequality over time concentrated over the 1980s. For example, the standard deviation of hourly wages for all jobs increased from a little under 0.50 in 1977 to around 0.60 in the early 1990s, before going down a bit in the mid-1990s. These changes are very similar to what has been documented using larger data sets such as the CPS or the U.S. Census.[26] More interestingly, the standard deviation in performance-pay jobs generally increases faster than in non-performance-pay jobs. This suggests that performance-pay jobs are closely linked to the growth of wage inequality since 1) inequality grew faster in performance-pay jobs, and 2) the growing incidence of performance-pay jobs means that an increasingly large fraction of workers are employed in this more unequal sector.

The main decomposition results are presented in Table VI and in Figures VI and VII. As before, the results are weighted using the PSID sample weights. Counterfactual distributions are obtained by multiplying the reweighting factor $\psi(X)$ by the PSID sample weight. As in Section IV, we also use the number of times a job match is observed to adjust for the end-point problem.[27]

Table VI shows that 21 percent of the increase in the variance of log wages can be accounted for by performance pay. More interestingly, the table also shows that most of the impact of performance pay is concentrated at the top-end of the wage distribution. In particular, performance pay accounts for only about 10 percent of the change in inequality at the bottom end of the distribution, as measured by the 50-10 gap (the difference between

[26] See Katz and Autor (1999) for an overview of the main trends in inequality based on several different data sources. We have also compared the inequality trends in the PSID to those in the March CPS when the same measure of the hourly wage (annual earnings divided by annual hours) and the same sample restrictions (heads only, private sector, age 18-65, not self-employed, and hourly wages between $1.50 and $100.00 in dollars of 1979) are used. In the PSID, the standard deviation of log wages increases from 0.501 in 1976-79 to 0.593 in 1990-98 (a 0.092 increase). The corresponding numbers in the March CPS are 0.508 in 1976-79 and 0.597 in 1994-98 (a 0.089 increase). The fact that the results are so similar in the two samples gives us great confidence in the representativeness of the PSID data.

[27] We perform this adjustment using yet another reweighting factor to adjust the distribution of the number of job-matches (in both the 1976-79 and 1990-93 periods) so that it is equal to the observed frequency distribution for the 1982-90 period.

the 50^{th} and the 10^{th} percentile of log wages).[28] By contrast, performance pay accounts for a large fraction -if not all- of the growth in inequality at the very top end of the distribution (the 99-90 or 99-75 gap).

This pattern is illustrated more dramatically in Figure VI that shows the difference between the actual and counterfactual wage distribution at each wage percentile.[29] The striking feature of the figure is that the effect of performance-pay jobs is concentrated at the top end of the wage distribution. It is also clear that the effect becomes larger in the early 1990s than in the late 1970s. Like Table VI, the figure shows that, as predicted by the model of Section II, the impact of performance pay is highly concentrated at the top-end of the distribution. Figure VII then compares the growth in wage inequality that would have prevailed with and without performance-pay jobs, by showing the change in real wages at each percentile in the actual (with performance-pay jobs) and counterfactual (without performance-pay jobs) wage distribution. Consistent with other studies, the figure shows that inequality grew faster in the top end than in the low end of the wage distribution.[30] The figure also shows that a very large fraction of the growth in wage inequality above the 80th percentile can be accounted for by performance-pay jobs.

VI.B. Variance Decomposition

While it would be tempting to conclude from these decompositions that the growth in performance pay <u>explains</u> 21 percent of the growth in the variance of wages, and most of the increase in inequality above the 80th percentile, the conclusion is too strong for several reasons. For instance, the "effect" of performance pay documented above depends both on the fraction of workers in performance-pay jobs, and on the relative effect of performance pay on the wage structure. The results reported in Table VI can either reflect the impact of changes in the fraction of performance-pay jobs, or simply that the inequality enhancing effect of performance-pay jobs has increased over time. As discussed in Section II, the two possible channels have very different implications for the role of performance pay in changes in wage inequality.

[28] The percentiles used to compute the measure of wage dispersion in Table VI are smoothed using a simple moving average to remove some sampling noise.

[29] Both Figures 6 and 7 are smoothed using local linear regressions.

[30] Though the difference between the evolution of top- and low-end inequality is particularly striking after the late 1980s (e.g. Autor, Katz, and Kearney (2006)), Table VI shows that the 90-50 gap expanded much more than the 50-10 gap (0.30 vs. 0.14) over our sample period.

A simple way of clarifying these issues is to compute a variance decomposition of the type that has been used in the literature on unions and wage inequality. Consider a simplified version of the wage equations for performance-pay and non-performance-pay jobs:

$$w^p = xb^p + e^p, \text{ and } w^n = xb^n + e^n.$$

The overall variance of wages across all workers can be written as $var(w) = V_1 + V_2 + V_3$, where

$$\begin{aligned} V_1 &= \overline{PPJ} \cdot var(xb^p|PPJ = 1) + (1 - \overline{PPJ}) \cdot var(xb^n|PPJ = 0), \\ V_2 &= \overline{PPJ} \cdot var(e^p|PPJ = 1) + (1 - \overline{PPJ}) \cdot var(e^n|PPJ = 0), \\ V_3 &= \overline{PPJ} \cdot (1 - \overline{PPJ}) \cdot \Delta^2, \end{aligned}$$

and where $\overline{PPJ}$ is the fraction of workers on performance-pay jobs, while $\Delta = E[x\beta^p + e^p|PPJ = 1] - E[x\beta^n + e^n|PPJ = 0]$ is the (raw) wage gap between the two types of jobs. The variance component V_1 captures how higher returns to observables among performance-pay workers contributes to wage inequality, while the variance component V_2 does the same for unobservables. The between-group component, or "wage gap" term, V_3, captures the fact that a positive wage gap between performance-pay and non-performance-pay jobs also tends to increase wage inequality.

Columns 1 and 4 of Table VII show the various components of the overall variance of wages in 1976-79 and 1990-93, respectively.[31] Columns 2 and 5 then show the counterfactual variance that would have prevailed if all workers had been paid according to the wage structure observed for non-performance-pay jobs. We do so by replacing the various components of the variance decomposition pertaining to performance-pay workers by the counterfactual components that would have prevailed if these workers had not been paid for performance.[32]

[31] We estimate the variance of the components xb and e by running standard regressions on the same variables used in Table IV, plus the additional interaction terms used to estimate the reweighting probits (four dummies in education interacted with experience, etc.).

[32] Following DiNardo and Lemieux (1997), we do so using a minor modification of the reweighting procedure described above. We need to replace $var(xb^p|PPJ = 1)$ with $var(xb^n|PPJ = 1)$, $var(e^p|PPJ = 1)$ with $var(e^n|PPJ = 1)$, and $E[xb^p + e^p|PPJ = 1]$ with $E[xb^n + e^n|PPJ = 1]$ in the definition of the wage gap component, Δ. To do so, we re-weight the non-performance-pay workers using the reweighting factor $\psi(x)/(1 - \psi(x))$ to get the distribution of wages that would have prevailed among performance-pay workers had they been paid like non-performance-pay workers. The counterfactual term $E[xb^n + e^n|PPJ = 1]$ is the mean of the resulting wage distribution, while $var(xb^n|PPJ = 1)$ and $var(e^n|PPJ = 1)$ are the explained and unexplained variances in a regression of wages on x in that counterfactual sample.

Columns 3 and 6 show the "effect" of performance pay by just taking the difference between the two other columns.

As indicated at the bottom of the table, performance pay accounts for 0.0290 of the 0.1361 growth in the variance of wages between 1976-79 and 1990-93. Most of the effect is linked to the impact of performance pay on observable determinants of wages. That component (row 3) increases by 0.0152 (from 0.0093 to 0.0245) over time, which represents over half of the total effect. The wage gap term (row 7) accounts for most of the remaining effect, while differences in the variance of the error terms (row 6) play a more modest role. These findings are consistent with the results in Table IV and V. Table IV shows that the return to education is higher in performance-pay jobs, and that this gap has increased over time. By contrast, while Table V also shows that the return to the worker-component is higher in performance-pay jobs, this is being offset by a larger idiosyncratic variance and a larger variance of the job-match term for non-performance-pay workers. On balance, performance pay does not have a large impact on the variance of the overall residual component (the sum of the three sub-components reported in Table V).

VI.C. Interpreting the Decomposition Results

The results reported in Table VII also enable us to answer, at least in part, the question raised in Section II about how the estimated "effect" of performance pay on inequality should be interpreted. Under a first scenario illustrated in Figure I, the growth in performance pay is just due to an exogenous decrease in the cost of implementing performance-pay schemes. Under the alternative scenario explored in Figure II, performance pay grows as a result of some underlying SBTC that induces more employers to use performance pay, and also increases the return to skill in performance-pay relative to non-performance-pay jobs.

If the first scenario was correct, then the contribution of performance pay to the growth in inequality should all be due to the increase in the share of workers paid more unequal wages (performance pay). This is inconsistent, however, with the fact that the effect of performance pay on the variance of wages of performance-pay workers linked to observables more than doubled from 0.0246 in 1976-79 to 0.0529 in 1990-93 (columns 3 and 6 of row 1 of Table VII). So even if the fraction of performance-pay workers had remained at 38 percent (1976-79 level) over time, the variance contribution would have still increased from 0.0093 (.38 times 0.0246) to 0.0201 (.38 times 0.0529). This represents over two thirds of the 0.0152 increase in the variance contribution (from 0.0093 to 0.0245) linked to observables shown in row 3.

In other words, most of the "effect" of performance pay on wage inequality is due to the fact that returns to observable skills increased faster in performance-pay than non-performance-pay jobs.[33] This is consistent with the scenario of Figure II where an increase in returns to skills induces more firms to adopt performance pay, but inconsistent with the simple story based on declining monitoring costs in Figure I. Our preferred interpretation of the results is, thus, that performance pay provides a channel through which underlying changes in the relative productivities of different groups of workers get translated into higher inequality.

Of course, this interpretation still leaves a very important role for performance pay in recent changes in inequality. Irrespective of why performance pay has increased over time, our decomposition results still indicate that, absent performance pay, wage inequality would have increased substantially less, and much less in the upper end of the wage distribution. But much of these inequality changes would not have happened either absent other underlying changes in the relative demand for skilled labor. It is in this sense that our findings should not be interpreted as the causal effect of the growth in performance pay on wage inequality where all other factors, including the relative demand for skilled labor, are held constant. Although some of the effect of performance pay on wage inequality may be due to exogenous developments linked to a decline in the cost of monitoring and information processing, our evidence suggest that this cannot account for most of the measured effect.

VII. Conclusion

An increasing proportion of jobs in the U.S. labor market include a performance-pay component in addition to regular wages and salaries. In this paper, we look at the connection between the growth of performance pay and wage inequality. The basic premise is that, relative to traditional (fixed-wage) jobs, wages on performance-pay jobs are more sensitive to productive characteristics of workers, and less sensitive to job characteristics. We develop a simple model to illustrate this point and derive several testable implications. Consist-

[33] The wage gap effect can also be linked to this phenomena. Performance-pay workers are more skilled than non-performance-pay workers, so an increase in the return to skills results in a larger between-group gap and in more inequality.

ent with predictions, we show that compensation in performance-pay jobs is more closely tied to both observed (by the econometrician) and unobserved productive characteristics of workers. As a consequence, wages are less equally distributed among performance-pay than non-performance-pay workers.

Building on these results we show that, in the absence of performance pay, the variance of males' wages would have increased by 21 percent less than it did between 1976-79 and 1990-93. Interestingly, most of the impact of performance pay on the growth in inequality is concentrated at the top end of the distribution. We find that inequality above the 80th percentile would have increased much more slowly in the absence of performance pay. This is a significant finding since most of the recent growth in wage inequality has been concentrated in this part of the wage distribution (Autor, Katz, and Kearney (2006)).

Our results also suggest that the growth in performance pay should not be thought of as an exogenous inequality enhancing change in the labor market that is unrelated to other labor market developments. In particular, the fact that the returns to skill increased faster in performance-pay than non-performance-pay jobs suggests that the growth in performance pay is, at least in part, an endogenous response by firms and workers to other underlying labor market developments. This is consistent with the view that performance pay provides a channel through which underlying changes in the relative productivities of different groups of workers get translated into higher inequality.

Going beyond the issue of wage inequality, this paper suggests that performance pay is not merely a way of packaging pay, but is also an integral part of production that can enhance the quality or worker-firm matches. In the absence of performance pay, workers and firms have to engage in costly search before workers with specific talents and abilities eventually get matched to the right job in the right firm. This has important consequences for the functioning of labor markets. For example, Shimer (2005) has shown that costly search may explain inter-industry wage differentials and why labor markets in the long run may fail to be perfectly competitive.[34] We conjecture that future research will find that performance pay systems also have a profound effect on wage dynamics, career concerns, and the overall efficiency of competitive labor markets.

[34]Efficiency wage models have often been cited as another potential explanation for inter-industry wages differences. Even in this case, as MacLeod and Malcomson (1988) show, sorting of workers into ability groups is not instantaneous, but can occur slowly over time.

University of British Columbia and NBER
Columbia University and NBER
McGill University

References

Abowd, John M., and David Card, "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57 (1989), 447–480.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney, " Rising Wage Inequality: The Role of Composition and Prices," NBER Working Paper No. 11628, 2005.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney, "The Polarization of the U.S. Labor Market," *American Economic Review*, 96 (2006), 189–194.

Baker, Michael, " Growth Rate Heterogeneity and the Covariance Structure of Life Cycle Earnings," *Journal of Labor Economics*, 15 (1997), 338–75.

Bebchuk, Lucian, and Jesse Fried, *Pay without Performance: The Unfulfilled Promise of Executive Compensation* (Cambridge, Massachusetts: Harvard University Press, 2004).

Bertrand, Marianne, and Sendhil Mullainathan, "Are CEOs Rewarded for Luck? The Ones without Principals Are," *Quarterly Journal of Economics*, 116 (2001), 901–932.

Bishop, John, "The Recognition and Reward of Employee Performance," *Journal of Labor Economics* , 5 (1987), S36–S56.

Brown, Charles, "Firm's Choice of Method of Pay," *Industrial and Labour Relations Review*, 43 (1990), 165S–182S.

Brown, James N., and Audrey Light, "Interpreting Panel Data on Job Tenure," *Journal of Labor Economics*, 10 (1992), 219–57.

Card, David, "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis," *Econometrica*, 64 (1996), 957–979.

Card, David, Thomas Lemieux, and W. Craig Riddell, "Unions and Wage Inequality," *Journal of Labor Research*, 25 (2004), 519–62.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64 (1996), 1001–1044.

DiNardo, John, and Thomas Lemieux, "Diverging Male Wage Inequality in the United States and Canada, 1981-1988: Do Institutions Explain the Difference?" *Industrial and Labor Relations Review*, 50 (1997), 629–51.

Freeman, Richard B., "How Much Has Deunionization Contributed to the Rise of Male Earnings Inequality?" in *Uneven Tides: Rising Income Inequality in America*, Sheldon Danziger and Peter Gottschalk, eds. (New York, NY: Russell Sage, 1993).

Frydman, Carola, and Raven Saks, "Executive Compensation: A New View from a Long-term Perspective, 1936-2005," Federal Reserve Board Technical Report No. 2007-35, 2007.

Gabaix, Xavier, and Augustin Landier, "Why has CEO Pay Increased so Much?" *Quarterly Journal of Economics*, 123 (2008), 49–100.

Gibbons, Robert, "Incentives and Careers in Organizations," in *Advances in Economics and Econometrics: Theory and Applications*, David M. Kreps and Kenneth F. Wallis, eds. (Cambridge, UK: Cambridge University Press, 1997).

Gibbons, Robert, Lawrence F. Katz, Thomas Lemieux, and Daniel Parent, "Comparative Advantage, Learning, and Sectoral Wage Determination," *Journal of Labor Economics*, 23 (2005), 681–723.

Goldin, Claudia, "Monitoring Costs and Occupational Segregation by Sex: A Historical Analysis," *Journal of Labor Economics*, 4 (1986), 1–27.

Harris, Milton, and Bengt Holmström, "A Theory of Wage Dynamics," *Review of Economic Studies*, 49 (1982), 315–333.

Jovanovic, Boyan, "Job Matching and the Theory of Turnover," *Journal of Political Economy* , 87 (1979), 972–90.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce, "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101 (1993), 410–442.

Katz, Lawrence F., and David H. Autor, "Changes in the Wage Structure and Earnings Inequality," in *Handbook of Labor Economics, Volume 3A*, Orley Ashenfelter and

David Card, eds. (Amsterdam: Elsevier North Holland, 1999)

Katz, Lawrence F., and Kevin M. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics* , 107 (1992), 35–78.

Kerr, Stephen, "On the Folly of Rewarding A, while Hoping for B," *Academy of Management Journal*, 18 (1975), 769–783.

Lawler III, Edward E., "Pay Practices in Fortune 1000 Corporations," *World at Work Journal*, 12 (2003), 45–54.

Lazear, Edward P., "Salaries and Piece Rates," *Journal of Business*, 59 (1986), 405–431.

Lazear, Edward P., "Performance Pay and Productivity," *American Economic Review*, 90 (2000), 1346–1361.

Lemieux, Thomas, "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review*, 96 (2006), 461–498.

Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent, "Supplemental Material to 'Performance Pay and Wage Inequality'," *http link on the QJE web site*, 2008.

MacLeod, W. Bentley, and Daniel Parent, "Job Characteristics and the Form of Compensation," *Research in Labor Economics*, 18(1999), 177–242.

MacLeod, W. Bentley, and James M. Malcomson, "Reputation and Hierarchy in Dynamic Models of Employment, " *Journal of Political Economy* , 96 (1988), 832–854.

Parent, Daniel, "Methods of Pay and Earnings: A Longitudinal Analysis," *Industrial and Labor Relations Review*, 53 (1999), 71–86.

Parent, Daniel, "Matching, Human Capital, and the Covariance Structure of Earnings," *Labour Economics*, 9 (2002), 375–404.

Piketty, Thomas, and Emmanuel Saez, "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, 118 (2003), 1–39.

Shimer, Robert, "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," *Journal of Political Economy*, 113 (2005), 996–1025.

Table I
Summary Statistics: Panel Study of Income Dynamics 1976-1998

	Non-performance-pay jobs [1]	Performance-pay jobs [2]
Average hourly earnings ($79)	8.38	10.86
Education	12.52	13.39
Potential experience	19.74	19.61
Employer tenure	7.62	9.25
Married	0.72	0.77
Unionized	0.28	0.14
Non white	0.13	0.09
Paid by the hour	0.66	0.31
Paid a salary	0.32	0.51
Annual hours worked	2,122.53	2,286.47
Number of workers (total:3,053)	2,616	1,271
Number of job matches (total: 7,442)	5,657	1,785
Number of observations (total: 26,146)	16,466	9,680

Notes: The sample consists of male household heads aged 18-65 working in private sector, wage and salary jobs. All figures in the table represent sample means. Education, potential experience, and employer tenure are measured in years. Potential experience is defined as age minus education minus 6. Performance-pay jobs are employment relationships in which part of the worker's total compensation includes a variable pay component (bonus, commission, piece rate). Any worker who reports overtime pay is considered to be in a non-performance-pay job. Workers are considered unionized if they are covered by a collective bargaining agreement.

Table II
Changes in the Incidence of Performance Pay between 1976-79 and 1990-93:
Robustness to Different Definitions and Contribution of Various Factors

	Minimum frequency of actual payments of performance pay			Received PP this year	Bonus only
	Any	1/5	1/2		
	[1]	[2]	[3]	[4]	[5]
1. Incidence in 1976-79	37.56	20.79	9.41	11.95	30.24
Change between 1976-79 and 1990-93					
2. Unadjusted change	12.92	6.78	1.51	4.45	12.36
3. Adjusted for the number of times a job match is observed	7.06	5.74	3.52	4.63	6.31
4. Row 3 plus adjustments for characteristics in row 5-9	4.57	3.93	3.00	3.70	4.79
Contribution of changes in characteristics (other than the number of times the job match is observed)					
5. Total (row 3 minus row 4)	2.49	1.80	0.52	0.93	1.52
6. Unions	1.44	0.83	0.18	0.37	1.28
7. Occupation	0.70	0.64	0.34	0.37	0.60
8. Industry	0.53	0.72	0.42	0.44	0.10
9. Other factors	-0.18	-0.40	-0.42	-0.25	-0.45

Note: All the adjustments and contributions of characteristics are computed by estimating linear probability models with a full set of dummies for periods (1976-1979, 1980-1984, 1985-1989, 1990-1993, and 1994-1998) and the number of times a job match is observed (1 to 22), as well as dummies for industry, occupation, marital status, race, union status, a cubic function in potential experience and a quadratic function in job tenure. 26,146 observations used in all columns.

Table III
Regression Estimates of the Effect of Performance Pay on Log Average Hourly Earnings

Estimation method:	OLS		Fixed effects		
	[1]	[2]	[3]	[4]	[5]
Performance pay job	0.0873	0.0597	0.0400	0.0225	-
	(0.0152)	(0.0166)	(0.0117)	(0.0120)	
Performance pay received in current year	-	0.0794	-	0.0380	0.0462
		(0.0167)		(0.0084)	(0.0059)
Worker fixed effect	No	No	Yes	Yes	Yes
Job-match fixed effect	No	No	No	No	Yes

Notes. 26,146 observations. Standard errors (in parentheses) are adjusted for clustering at the job-match level. All specifications also include a full set of industry (10), occupations (8), and year (22) dummies, a cubic in potential experience, a quadratic in job tenure, years of completed schooling, calendar year average of the unemployment rate in the county of residence, and dummies for being married, nonwhite, and for union status. The "performance-pay job dummy" indicates if either a bonus or commission/piece rate earnings are received at any time during the employment relationship; the "performance pay received in current year" dummy indicates if a bonus or commissions/piece rates earnings are received in the current year.

Table IV
Skills Related Wage Differentials and Performance-Pay (PP) Jobs

Sample:	PP jobs	Non-PP jobs	All jobs			
Estimation method:	OLS	OLS	OLS	FE	OLS	FE
	[1]	[2]	[3]	[4]	[5]	[6]
Performance-pay job dummy	-	-	-0.4526	-0.2061	-0.2406	0.1414
			(0.1019)	(0.0723)	(0.1251)	(0.0998)
Years of education	0.0929	0.0665	0.0637	0.0167	0.0584	0.0040
	(0.0071)	(0.0039)	(0.0040)	(0.0091)	(0.0047)	(0.0096)
Education X performance-pay job	-	-	0.0365	0.0169	0.0217	-0.0079
			(0.0071)	(0.0048)	(0.0092)	(0.0071)
Education X 1990-93	-	-	-	-	0.0161	0.0222
					(0.0085)	(0.0056)
Education X performance-pay job	-	-	-	-	0.0190	0.0280
X 1990-93					(0.0137)	(0.0089)
Potential experience:	0.4259	0.2882	0.3010	0.4545	0.3002	0.4231
(effect at 20 years)	(0.0535)	(0.0288)	(0.0294)	(0.1258)	(0.0294)	(0.1256)
Experience X performance-pay job:	-	-	0.1162	0.0149	0.1018	-0.0278
			(0.0584)	(0.0501)	(0.0581)	(0.0509)
Tenure:	0.1670	0.2197	0.2262	0.1158	0.2271	0.1191
(effect at ten years)	(0.0268)	(0.0154)	(0.0154)	(0.0129)	(0.0154)	(0.0129)
Tenure X performance-pay job:	-		-0.0666	0.0278	-0.0677	0.0196
			(0.0301)	(0.0237)	(0.0303)	(0.0239)
Number of observations	9,680	16,466	26,146	26,146	26,146	26,146

Notes: Standard errors (in parentheses) are adjusted for clustering at the job-match level. All specifications also include a full set of industry (10), occupations (8), and year (22) dummies, a cubic in potential experience, a quadratic in job tenure, years years of completed schooling, calendar year average of the unemployment rate in the county of residence, and dummies for being married, race, and for union status. The reported effects of potential experience (at 20 years) and tenure (at 10 years) are the predicted levels computed using the estimated polynomial models. The models in columns 3 to 6 include interactions between the performance-pay dummy and education, a cubic in potential experience, and a quadratic in tenure. The models in columns 5 and 6 include a full set of interactions between period dummies for 1980-1984, 1985-1989, 1990-1993, and 1994-1998, education and the performance-pay job dummy, but only the estimates for 1990-1993 are reported. The acronym FE refers to the fixed-effect method (worker fixed-effect).

Table V
Variance Component Models by Type of Job

Panel A: full sample

	Performance-pay jobs			Non-performance-pay jobs		
Parameter	[1]	[2]	[3]	[4]	[5]	[6]
Variance of	0.102	0.102	0.082	0.068	0.057	0.047
worker component	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)	(0.002)
Factor loading: 1990-93	-	-	1.202	-	-	1.173
relative to 1976-79			(0.033)			(0.034)
Variance of job-	-	0.004	0.004	-	0.018	0.017
match component		(0.003)	(0.003)		(0.002)	(0.002)
Variance of	0.083	0.082	0.096	0.098	0.091	0.093
idiosyncratic error	(0.003)	(0.003)	(0.008)	(0.002)	(0.002)	(0.004)
Change in variance,	-	-	-0.011	-	-	0.024
1976-79 to 1990-93			(0.011)			(0.006)
Number of workers	1,271	1,271	1,271	2,616	2,616	2,616
Number of cross-products	64,486	64,486	64,486	99,554	99,554	99,554

Panel B: workers who worked in both types of jobs

	Performance pay jobs			Non performance pay jobs		
Parameter	[1]	[2]	[3]	[4]	[5]	[6]
Variance of	0.104	0.102	0.067	0.065	0.053	0.036
worker component	(0.002)	(0.004)	(0.006)	(0.002)	(0.002)	(0.004)
Factor loading: 1990-93	-	-	1.312	-	-	1.309
relative to 1976-79			(0.061)			(0.105)
Variance of job-	-	0.002	0.000	-	0.026	0.023
match component		(0.004)	0.004		(0.004)	(0.004)
Variance of	0.085	0.085	0.114	0.108	0.094	0.082
idiosyncratic error	(0.004)	(0.004)	(0.012)	(0.004)	(0.004)	(0.009)
Change in variance,	-	-	-0.027	-	-	0.035
1976-79 to 1990-93			(0.015)			(0.013)
Number of workers	834	834	834	834	834	834
Number of cross-products	32,476	32,476	32,476	52,073	52,073	52,073

Note: Standard errors in parentheses. Models in columns 3 and 6 allow the variance of the idiosyncratic errors and the factor loadings on the worker component to vary across the 1976-1979, 1980-1984, 1985-1989, 1990-1993, and 1994-1998 periods, while models in columns 1, 2, 4, and 5 do not. These equally weighted covariance structure models are fit to the cross-products of the residuals of an OLS regression of log wages on the same set of covariates described in Table IV. Note that the factor loadings in columns 3 and 6 are normalized to 1 in the base period (1976-1979), so that the changes in factor loadings can be interpreted as the percentage changes in the return to the worker component.

Table VI
Reweighting Estimates of the Effect of Performance-Pay (PP) Jobs on Measures of the Distribution of Log Wages

	1976-79			1990-93			Percentage of the change explained by PP jobs
	Actual dispersion	Dispersion w/o PP jobs	Effect of PP jobs	Actual dispersion	Dispersion w/o PP jobs	Effect of PP jobs	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Variance	0.229	0.209	0.020	0.365	0.317	0.049	21.2
Percentile gaps:							
90-10	1.187	1.145	0.042	1.559	1.472	0.086	12.0
50-10	0.653	0.649	0.003	0.790	0.772	0.017	10.2
90-50	0.534	0.496	0.039	0.769	0.700	0.069	13.1
99-90	0.566	0.469	0.097	0.715	0.465	0.250	102.7
75-50	0.285	0.258	0.027	0.373	0.331	0.042	16.9
99-75	0.852	0.727	0.124	1.088	0.796	0.292	70.9

Notes: The counterfactual measures of wage dispersion reported in columns 2 and 5 are computed by reweighting the sample of workers in non-performance-pay jobs (see text for details). The effect of performance-pay reported in columns 3 and 6 is the difference between columns 1 and 2 and columns 4 and 5, respectively. The number of observations used is 4,719 and 4,913 for 1976-1979 and 1990-1993, respectively. The percentile gaps are the differences between the two corresponding percentiles of the log wage distribution. These percentiles are smoothed using a moving average (window of +- 2 percentiles) to reduce sampling noise. Column 7 is the ratio (in percentage terms) of the difference between columns 6 and 3 over the difference between column 4 and 1.

Table VII
The Contribution of Performance-Pay (PP) Jobs to the Variance of Log Hourly Earnings

	1976-1979			1990-1993		
	Actual variance	Variance w/o PP jobs	PP jobs effect	Actual variance	Variance w/o PP jobs	PP jobs effect
	[1]	[2]	[3]	[4]	[5]	[6]
Within-group variance due to observables						
1. Var(XB | PP Jobs=1)	0.1280	0.1034	0.0246	0.2037	0.1507	0.0529
2. Var(XB | PP jobs=0)	0.0922	0.0922	0.0000	0.1420	0.1420	0.0000
3. Total variance (%(PP jobs)*row 1 + (1-%(PP jobs))*row 2)	0.1057	0.0964	0.0093	0.1706	0.1460	0.0245
Within-group variance due to unobservables						
4. Var(e | PP jobs=1)	0.1220	0.1085	0.0135	0.1773	0.1583	0.0189
5. Var(e | PP jobs=0)	0.1139	0.1139	0.0000	0.1690	0.1690	0.0000
6. Total variance (%(PP jobs)*row 4 + (1-%(PP jobs))*row 5)	0.1170	0.1119	0.0051	0.1728	0.1640	0.0088
Between-group variance (Wage Gap Effect)						
7. %(PP jobs)*(1-%(PP jobs))*Δ^2	0.0062	0.0009	0.0054	0.0217	0.0062	0.0155
Overall variance of wages						
8. Var(XB + e): (row 3 + row 6 + row 7)	0.2290	0.2091	0.0198	0.3650	0.3163	0.0488
Fraction of performance-pay jobs (%(PP jobs))		0.3783			0.4632	
Change in overall variance (col. 4 – col. 1):	0.1361					
Change in performance pay job effect (col. 6 – col. 3):	0.0290					
Share of performance pay job effect:	21.28%					

Notes. Computations for the counterfactual variances (columns 2 and 5) done by reweighting to produce a counterfactual distribution for performance-pay workers. Δ is the difference in mean wages between performance-pay and non-performance-pay workers. The samples for 1976-1979 and 1990-1993 are also adjusted by reweighting so that the distribution of the number of job matches observed is the same as in 1982-1990. The wage regression estimated to divide wages into an explained (XB) and an unexplained (e) component uses a more flexible specification in the explanatory variables listed in the notes to Tables III and IV. Relative to these specifications, we add a set of four education dummies that we also interact with potential experience (linear term), union status, and the race dummy. We also add a cubic in tenure, a dummy for full-time/full-year workers, and an interaction between potential experience and the race dummy. The probit used for reweighting also uses the same specification. See the text for more detail.

Table A.1
Incidence of Performance Pay Jobs by Industry and Occupation, PSID 1976-1998

Industry categories	Occupation categories (1 digit)								
(1 digit)	Professionals	Managers	Sales	Clerical	Craftsmen	Operatives	Laborers	Services	Total
Min.& durables	0.46	0.60	0.74	0.35	0.21	0.26	0.26	0.19	0.33
Non-durables	0.52	0.62	0.79	0.38	0.26	0.42	0.25	0.06	0.43
Transpo. & utils	0.23	0.52	0.82	0.28	0.27	0.37	0.44	0.37	0.35
Fin., insur., & real est.	0.75	0.72	0.87	0.38	0.29	0.04	0.22	0.33	0.65
Bus. & prof. serv.	0.41	0.57	0.73	0.39	0.46	0.40	0.18	0.25	0.43
Personal serv.	0.42	0.63	0.61	0.24	0.33	0.20	0.29	0.49	0.46
Whol-tr.& oth serv.	0.65	0.66	0.82	0.45	0.29	0.46	0.29	0.03	0.58
Retail trade	0.27	0.57	0.72	0.33	0.48	0.32	0.21	0.46	0.50
Construction	0.72	0.47	0.81	0.20	0.33	0.30	0.30	0.17	0.36
Agriculture & fishing	0.72	0.78	0.88	0.24	0.16	0.42	0.45	0.77	0.46
Total	0.45	0.59	0.78	0.34	0.30	0.33	0.31	0.34	0.42

Note: The (adjusted) proportions are computed over the sample of 26,146 observations used in Table I. The proportions are adjusted by running a linear probability model for the incidence of performance-pay jobs on a full set of dummies for the number of times a job match is observed, and computing the predicted incidence holding the distribution of the number of times the job match is observed at its average value for the 1982-90 period.

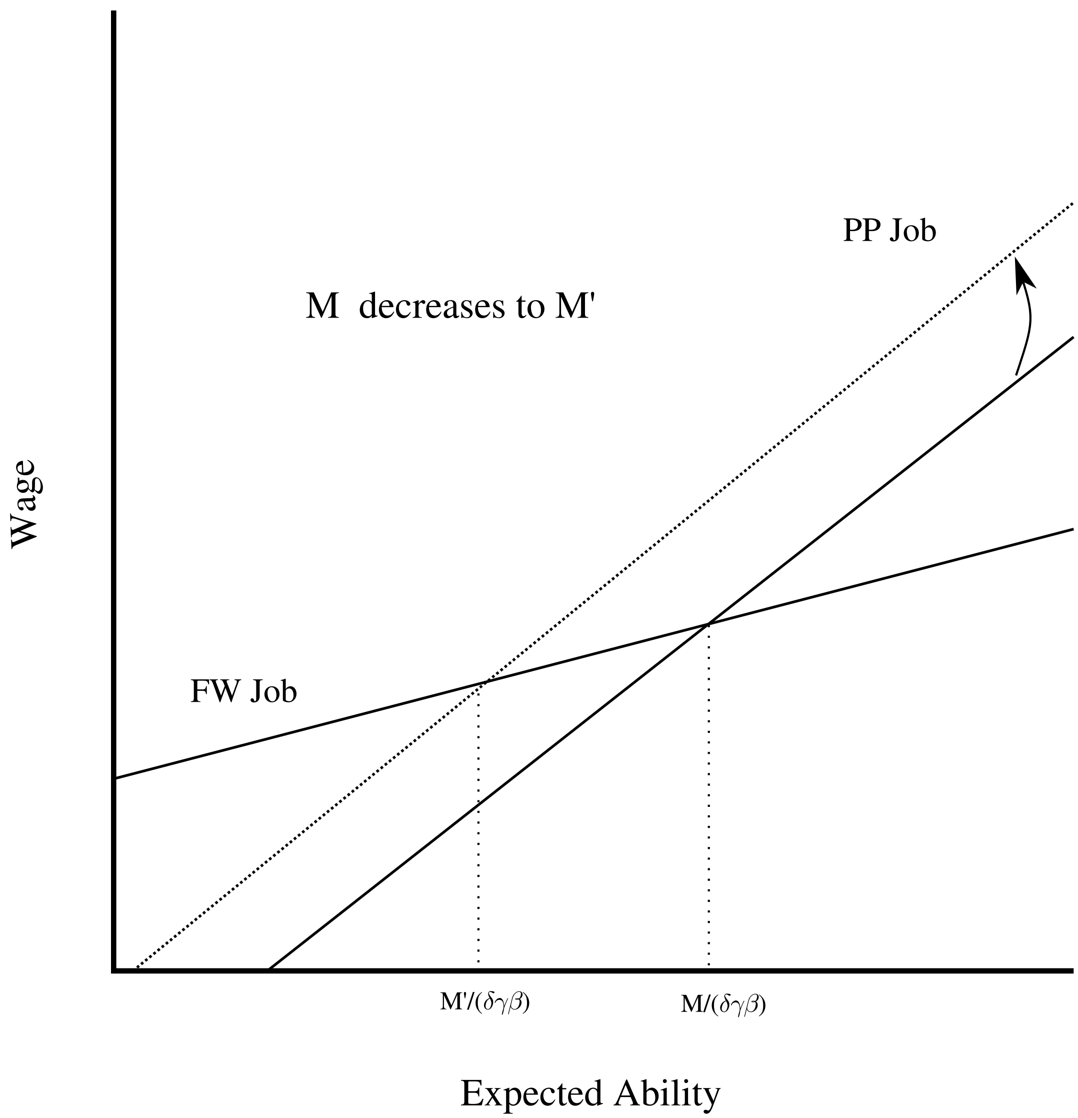

Figure I

Effect of Monitoring Cost Decrease

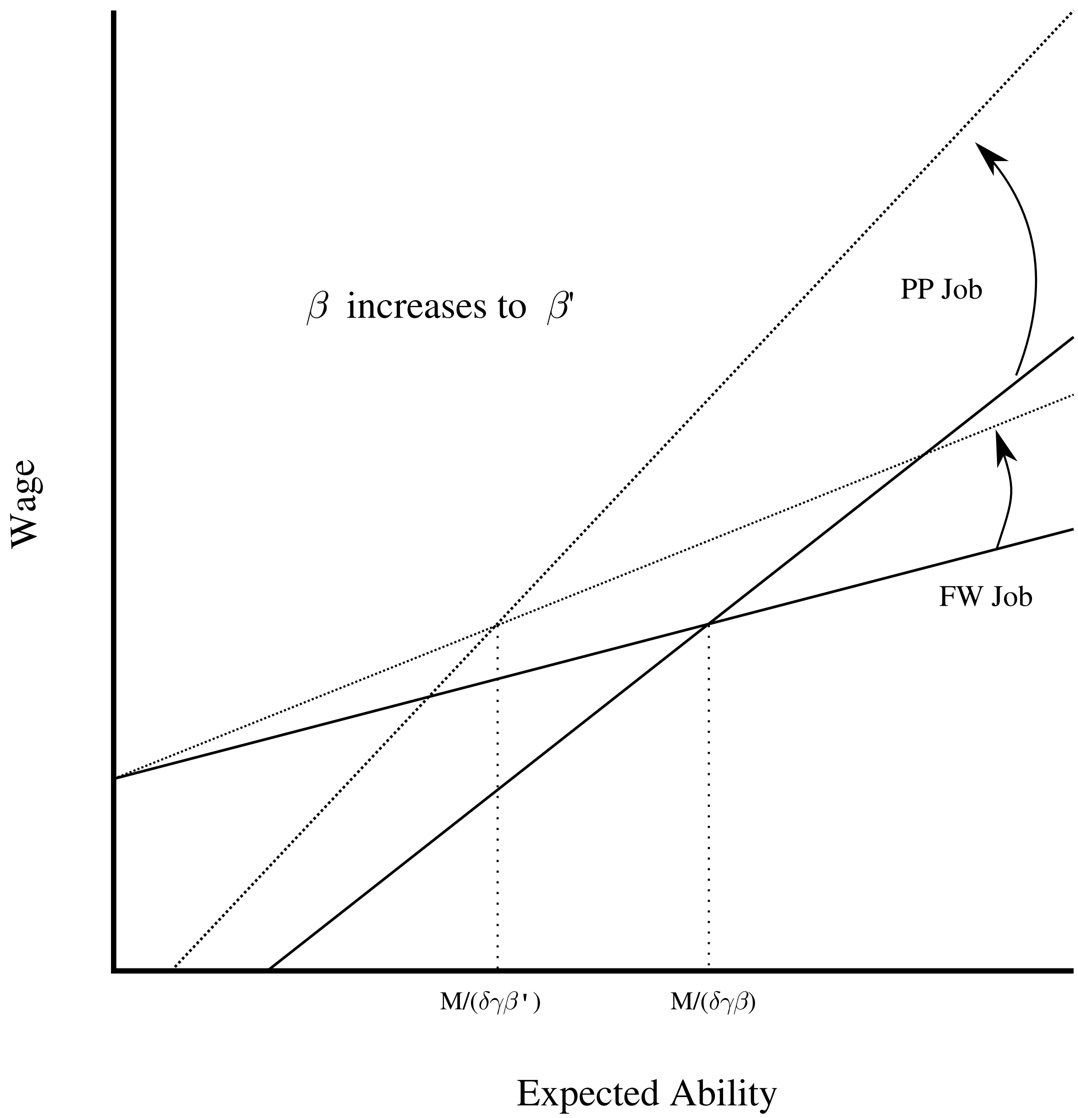

Figure II

Effect of SBTC

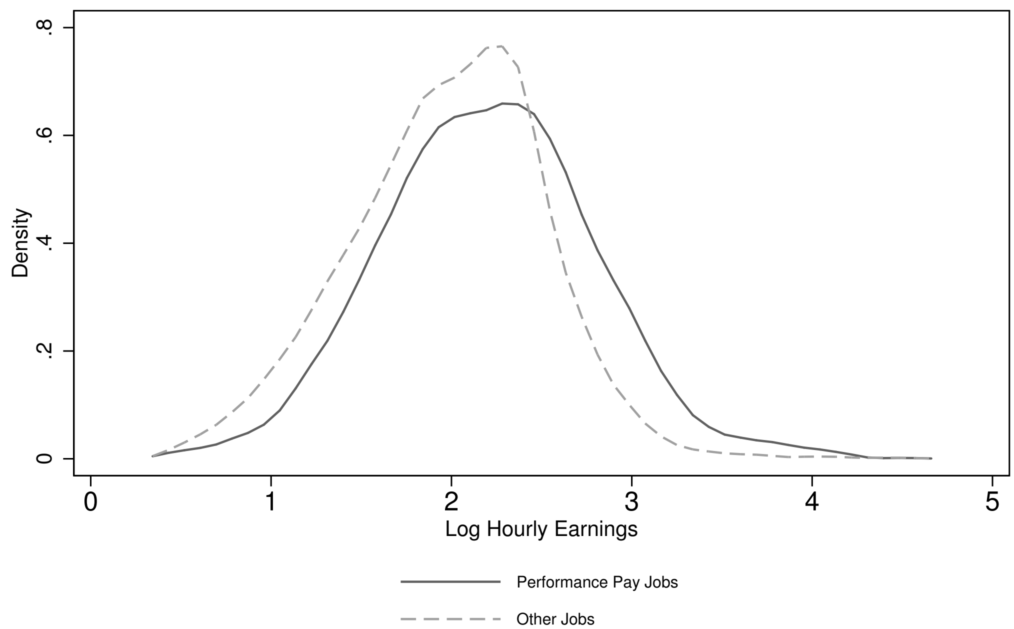

Figure III
Distribution of Log Hourly Earnings: PSID 1976–1998

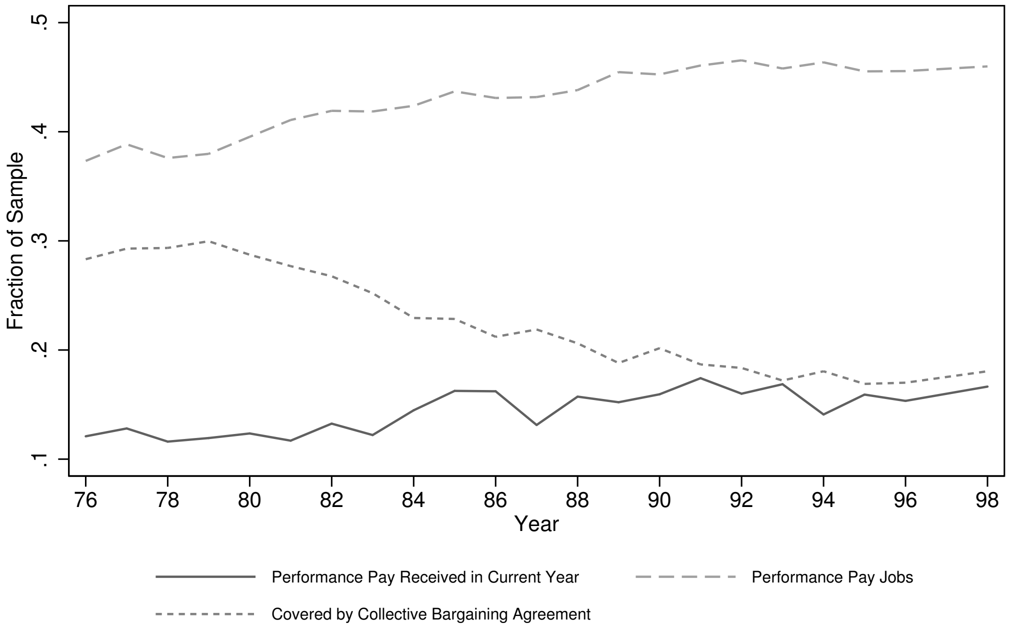

Figure IVa
Performance Pay Job Incidence: PSID 1976–1998

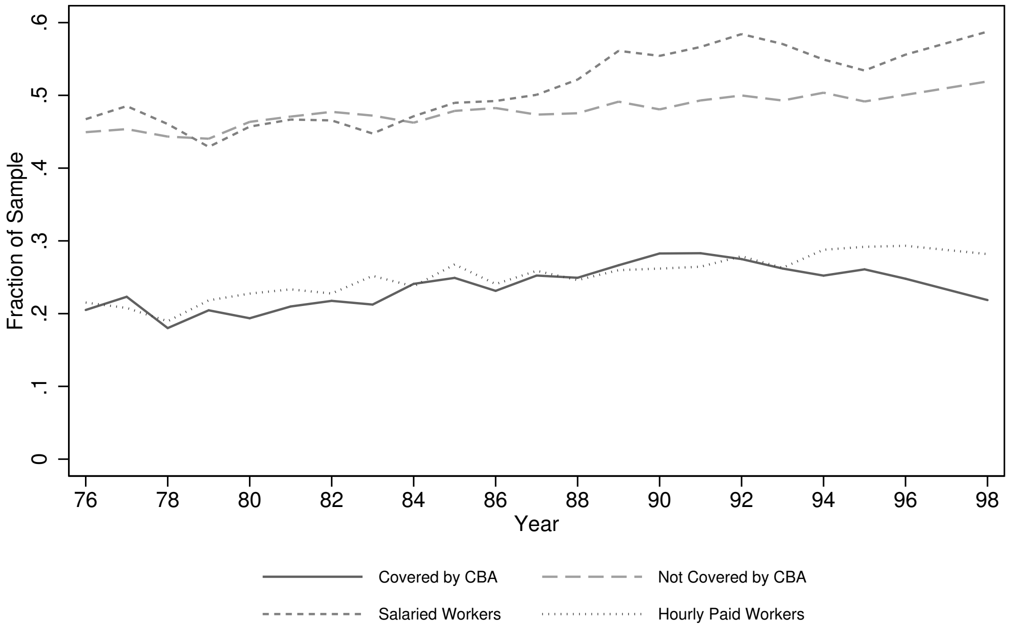

Figure IVb
Performance Pay Job Incidence By Types of Workers: PSID 1976−1998

Figure V
Wage Inequality: PSID 1976−1998

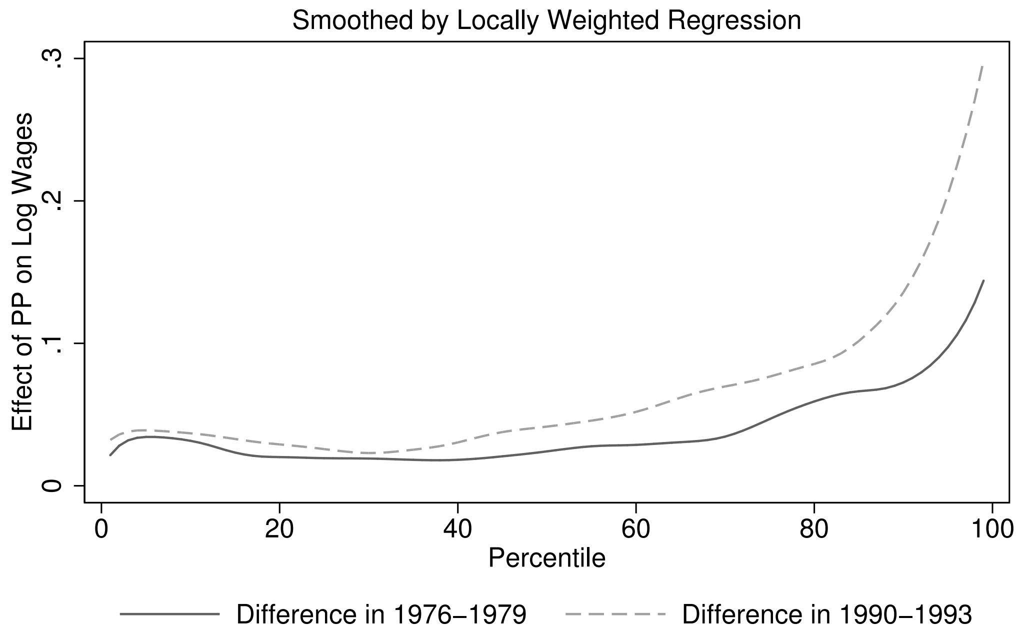

Figure VI
Effect of Performance Pay on Wages at each Percentile: PSID 1976−1998

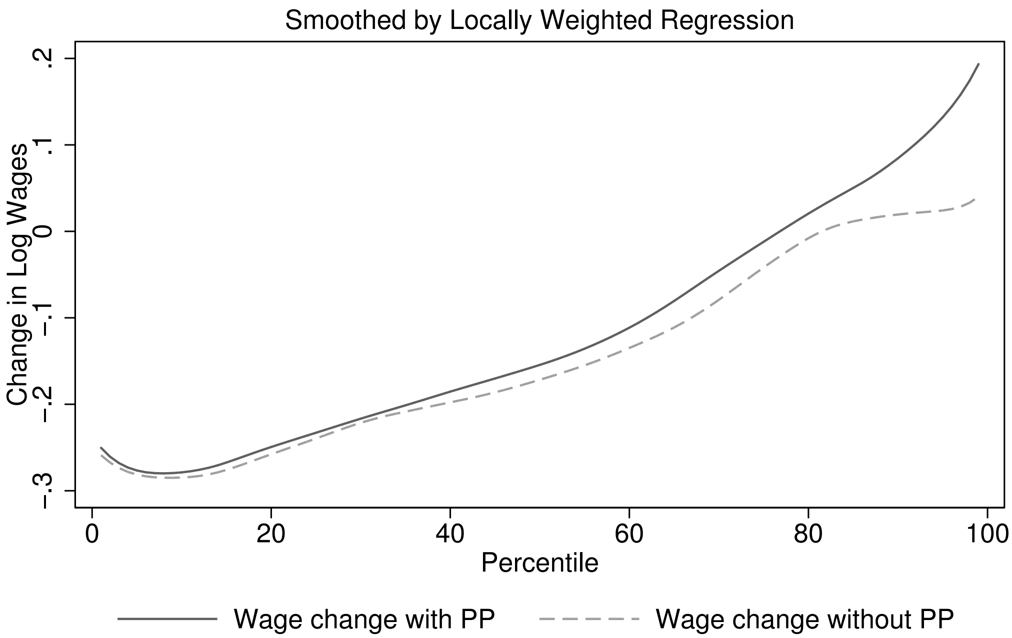

Figure VII

Change Over Time in Wages by Percentile with and without Performance Pay Jobs: PSID1976−1998