Supplementary Material to "Performance Pay and Wage Inequality"

Thomas Lemieux
W. Bentley MacLeod
Daniel Parent

June 2008

This document provides supplementary material to our paper "Performance Pay and Wage Inequality". We first provide in Appendix 1 a more detailed derivation of the model used in Section II of the paper. We then present in Appendix 2 a parametric measurement model used to adjust the estimated effect of performance pay on wages for the fact that shorter jobs are less likely to be classified as performance-pay jobs since actual performance-related payments are not observed each and every year. Appendix 3 shows some evidence of the robustness of our main results reported in Table IV using alternative specifications with the PSID data, as well as a separate analysis based on data from the NLSY. Finally, Appendix 4 provides some additional details on the identification and estimation of the variance components models presented in Section V.C of the paper. We also include two additional tables and one additional figure that we refer to in this document and in the paper.

Appendix 1: Performance Pay, Monitoring Costs and Skill Biased Technical Change

This appendix presents a simple model of compensation choice that illustrates how changes in monitoring costs and skill biased technical change affect inequality and the mix of job types. The model builds upon Lazear (1986)'s observation that the reason performance

pay is used is because at the time a worker is employed neither the worker nor the firm can perfectly observe ability. This can 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 increase overall productivity. However, the introduction of an effective performance-pay system is expensive and difficult (see Kerr (1975)).[1] Hence, one faces a trade-off between the cost of introducing such a system, and the benefits in terms of improved match quality.

The detailed predictions of our model depend upon the distribution of job and individual characteristics. Nevertheless, two robust predictions arise. First, a decrease in monitoring (or related information processing) costs always increases the probability that firms will use performance pay instead of fixed wages. Second, performance pay increases wage dispersion relative to a payment system based on straight wages.

More formally, consider a competitive labor market in which the workers obtain all the rents from any match. Hence, in contrast to union settings, in this case there will be no tension between the worker's and the employer's preferences over contract form. We will suppose that there is symmetric information, namely the worker and employer have the same information regarding the employee's characteristics. For simplicity, suppose that the observed characteristics of the individual, x_i, are a sufficient statistic describing the individual's expected ability. We assume the market to be in equilibrium and that the individual is *ex ante* indifferent between a number of jobs $j \in J = \left\{1, .., \bar{J}\right\}$, each with observed characteristics z_j.

When worker i approaches firm j, they engage in the following sequence of moves:

1. The worker observes y_j and the firm makes a take it or leave if offer of a fixed-wage or a performance-pay contract:

   (a) A fixed wage contract is of the form $c_{ij} = \{w_{ij}, \bar{e}_{ij}\}$, where the worker agrees to supply effort $\bar{e}_{ij}$ in exchange for a wage w_{ij}. It is assumed that this contract is enforceable, and performed as promised.

   (b) A performance wage contract $w_{ij}(y_{ij}) = y_{ij}$, where y_{ij} is output.

2. After accepting the contract the worker observes his individual ability and chooses "effort" e_{ij}.

[1] See also Baker (1992) and Holmström and Milgrom (1987).

3. Output $y_{ij} = k_j + \beta\gamma_j \bar{e}_{ij}$ under the fixed wage contract is produced, while $y_{ij} = k_j + \beta\gamma_j e_{ij} - M_j$ is produced under the performance pay contract, where:

   (a) k_j is output that is independent of effort.

   (b) $\gamma_j > 1$ is the marginal product of effort.

   (c) M_j is the fixed cost of putting in a system to monitor effort.

   (d) $\beta \geq 1$ parametrizes the skill bias in technology that may vary with time and worker characteristics.

   The worker is paid a wage w_{ij} according to the contract terms and gets utility:

$$U_{ij} = w_{ij} - exp\left(e_{ij} - \alpha_i\right),$$

   where ability is given by the latent variable, $\alpha_i \sim N\left(\hat{\alpha}_i, \sigma_i^2\right)$. It is assumed that the mean and variance are known conditional upon worker characteristics x_i: $\hat{\alpha}_i = E\left\{\alpha_i | x_i\right\}$and $\sigma_i^2 = var\left\{\alpha_i | x_i\right\}$.

We now address three questions. First, what contract form will a particular match choose? Second, how does compensation vary as a function of job characteristics? Third, what is the consequence of an exogenous change in the return to skills for the choice between using performance pay or a fixed wage contract?

Optimal Contract Choice

Consider a particular match. Under the assumption that the worker receives all the rent from a relationship, we can derive the optimal contract as follows. Under a fixed wage contact, $c_{ij} = \{w_{ij}, \bar{e}_{ij}\}$, the worker obtains:

$$\begin{aligned} U_i\left(c_{ij}\right) &= E\left\{k_j + \beta\gamma_j \bar{e}_{ij} - exp\left\{\bar{e}_{ij} - \alpha_i\right\}\right\} \\ &= k_j + \beta\gamma_j \bar{e}_{ij} - exp\left\{\bar{e}_{ij} - \hat{\alpha}_i + \sigma_i^2\right\}. \end{aligned}$$

From this the optimal effort is $\bar{e}_{ij}^* = log\left(\beta\gamma_j\right) + \hat{\alpha}_i + \sigma_i^2$, from which we obtain the optimal utility for a fixed-wage job in match ij:

$$U_{ij}^{FW} = k_j + \beta + \gamma_j log\left(\beta\gamma_j\right) + \beta\gamma_j\left(\hat{\alpha}_i - \sigma_i^2\right) - \beta\gamma_j,$$

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

where without loss of generality we let $m_j = k_j + \beta\gamma_j log(\beta\gamma_j)$, and hence the observed wage is given by:

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

Now consider the performance pay situation, where $w_{ij}(e) = k_j + \gamma_j e$, and the individual is able to choose effort after observing ability α_i. In this case the payoff is:

$$U_i\left(c_{ij}\right) = k_j + \beta\gamma_j e(\alpha_i) - exp\left\{e(\alpha_i) - \alpha_i\right\},$$

and optimal effort satisfies $e_{ij}^*(\alpha_i) = log\left(\beta\gamma_j\right) + \alpha_i$, and expected utility from a performance pay contract satisfies:

$$\begin{aligned} U_{ij}^{PP} &= E\left\{k_j + \beta\gamma_j\left(log\left(\beta\gamma_j\right) + \alpha_i\right) - exp\left\{log(\gamma_j)\right\} - M_j\right\} \\ &= m_j + \beta\gamma_j\hat{\alpha}_i - M_i - \beta\gamma_j. \end{aligned}$$

This is the expected utility before accepting the job. Under a performance-pay job the *observed wage* is given by:

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

while the *ex ante* expected wage is:

$$\hat{w}_{ij}^{PP} = m_j + \beta\gamma_j\hat{\alpha}_i - M_j,$$

Given these we now have:

Proposition 1 In a match between worker i and firm j a performance pay contract is used if and only if:

$$\hat{w}_{ij}^{PP} \geq w_{ij}^{FW}, \tag{1}$$

and hence if and only if:

$$\sigma_i^2 \geq \frac{M_j}{\beta\gamma_j}. \tag{2}$$

Equation (1) shows that in this model the expected wage in both types of contracts should be the same conditional on job characteristics j. As we shall see, performance pay

contracts do pay more, in part because of the selection of more productive workers into those jobs.

The second inequality (equation (2)) shows how workers of different ability levels self-select themselves into performance-pay jobs. This highlights the point made in Lazear (1986) that the benefit of performance pay lies in its ability to tailor the demands of the job to the individual. It is common for people to think of performance pay as a way to "extract more effort," and hence that it is likely to be preferred by more skilled individuals. Such reasoning does not explain the wide use of performance pay in some low skilled jobs, such as agricultural work or sales. In these occupations the cost of monitoring, M_j, is low while there is a great deal of variance in ability (in agricultural work often the whole family, including children, are involved). In these cases it is more efficient to use performance pay because it allows more efficient matching of skills to the job for both high and *low* ability individuals.

Conditional upon job and worker characteristics the variance of wages in performance pay jobs is higher:

$$var\left(w^{PP}|x_i, z_j\right) > var\left(w^{FW}|x_i, z_j\right).$$

Finally, an important question is understanding why performance pay is becoming more prevalent over time. One possibility is SBTC, another is a change in the cost of monitoring performance. For a given job match, an increase in β due to SBTC or a decrease in the cost of monitoring, M_j, may cause the firm to switch from a fixed-wage contract to a performance-pay contract. Moreover, conditional upon performance pay being offered, SBTC increases the variance of observed wages, while a change in monitoring costs has no effect.

The Market Conditional Upon Individual Characteristics

Consider now the implications of the model for equilibrium wages conditional upon x_i. Suppose that there are n workers, assigned to jobs $j \in J = \{1, .., n\}$. Here, "jobs" are best viewed as different combinations of industry and occupation that differ in terms of their marginal products γ_j (and M_j and k_j) for technological reasons. For this to be an equilibrium, it must be the case that the ex ante utility is the same in all jobs: $\hat{U}_i = U_{ij} = U_{ij'}$ for all $j, j' \in J$. Taking $\hat{U}_i$ as given, consider first fixed-wage jobs. For fixed-wage jobs, we have:

$$w_{ij}^{FW} = \hat{U}_i + \beta\gamma_j^{FW},$$

while for performance-pay jobs, we have

$$w_{ij}^{PP} = \hat{U}_i + \beta\gamma_j^{PP} + \beta\gamma_j^{PP}\left(\alpha_i - \hat{\alpha}_i\right),$$

where we use the superscripts FW and PP to indicate which jobs pay fixed wages and performance pay, respectively, in equilibrium. The job effects $\beta\gamma_j^k$, $k = FW, PP$, can be interpreted as compensating differentials linked to the fact that workers have to provide more effort on jobs with a higher marginal product γ_j^k. Since the marginal return to effort/skills in performance-pay jobs is higher than in fixed-wage jobs ($\gamma_j^{PP} > \gamma_j^{FW}$), the expected wage in performance-pay jobs is higher. As we observed above, holding utility fixed, this also implies that the expected wage is the same in both types of jobs, while the variance is higher in performance-pay jobs.

We can also get an effort effect. However, this occurs not because performance pay *causes* higher effort, but because of selection of jobs types. Suppose that the variance of the cost of monitoring is relatively small across jobs in J, and is approximately equal to M. In that case, we have that *conditional upon J,* a job is performance pay if and only if:

$$\gamma_j^{PP} \geq \frac{M}{\beta\sigma_i^2} > \gamma_j^{FW}.$$

From this we find that

$$\begin{aligned} E\left\{w_{ij}^{PP} - w_{ij}^{FW}|x_i\right\} &= E\left\{\beta\gamma_j|x_i, PP\right\} - E\left\{\beta\gamma_j|x_i, FW\right\} \\ &= \beta\left\{E\left\{\gamma_j|x_i, PP\right\} - E\left\{\gamma_j|x_i, FW\right\}\right\} \\ &> 0. \end{aligned}$$

In other words, the average pay of workers, conditional upon their ability, should be higher in performance-pay than in fixed-wage jobs. Moreover, the mean difference should increase when there is SBTC.

Notice that this result very much depends upon the relative homogeneity of monitoring costs. If jobs differ mainly due to monitoring costs, then the result may not hold. In particular, it implies that conditional upon job characteristics and worker characteristics, the difference between performance-pay and fixed-wage jobs should be greater than if one simply conditions upon workers' characteristics.

Effect of Job Characteristics on Wages

Given that the job effect differs across jobs in the equations above, and that it is interacted with unobserved ability $\alpha_i - \hat{\alpha}_i$ under performance pay, it is not clear, a priori, whether the variance of wages induced by the job effects is larger under performance-pay or fixed-wage contracts.

Intuitively, however, we expect the fixed-wage job effect $\beta\gamma_j^{FW}$ to explain more of the variance of wages than is the case for $\beta\gamma_j^{PP}$. For the sake of simplicity, up to now we have assumed that the choice of job j did not depend on unobserved ability. But relaxing this assumption under performance pay would further improve the quality of the match between workers and jobs as high unobservable ability workers would sort themselves into high marginal product jobs, as in Gibbons, Katz, Lemieux, and Parent (2005). As a result of this complementarity between the job effect and worker ability, a substantial part of the job effect $\beta\gamma_j^{PP}$ would be absorbed by the individual ability term α_i. Econometrically speaking, this means that much of these job effects would disappear after controlling for a worker-specific fixed effect.

Furthermore, these conventional job effects linked to industry and occupation characteristics do not capture more fundamental differences in what a job represents under fixed wages relative to performance pay. In our stylized model, each worker with different observed characteristics x_i gets a different fixed-wage job created for her. By contrast, workers with different observable characteristics all share the same broadly defined job under performance pay, but then tailor their effort on the job to both their observable and unobservable characteristics. But since there is a one–to–one mapping between observed characteristics x_i and the fixed-wage jobs, it is not possible to empirically distinguish the "effect" of these jobs on wages from the effect of x_i on wages.

In practice, however, it is unlikely that firms can tailor a fixed-wage job to the precise characteristics x_i of each worker. For example, in the presence of search costs, firms will still create fixed-wage jobs based on expectations regarding the characteristics of workers they plan to hire. Yet, workers who actually apply and are hired will not necessarily have the optimal characteristics for the job. As a result, workers with slightly different values of x may end up working on the same job for the same employer, and get paid the same wage for doing so. By contrast, wages remain unaffected by these search frictions in performance-pay jobs because workers can tailor their effort to both observable and unobservable job

characteristics.

Two interesting empirical implications emerge from this minor extension of our model. First, the same worker in industry-occupation j can be paid different wages for jobs with different employers under fixed wages, but there should be no such employer wage effect under performance pay. Second, the imperfect matching between jobs and observable characteristics under fixed wages provides an additional reason why returns to observable worker characteristics will be lower on fixed-wage than performance-pay jobs.

Appendix 2: Measurement Error Correction

Let PPJ^* be a dummy indicating whether a job is truly one that pays for performance. The probability that $PPJ^* = 1$ depends on observed characteristics X (education, occupation, etc.):

$$\Pr(PPJ^* = 1|X) = q(X).$$

Let P represent the probability that we observe a performance payment in a given time period, conditional on $PPJ^* = 1$, and let T represent the number of observations we have for a job match. It follows that:

$$\Pr(PPJ = 1|X, T) = q(X, T) = q(X)[1 - (1 - P)^T],$$

and

$$\Pr(PPJ^* = 1|PPJ = 1, X, T) = \Pr(PPJ^* = 1|PPJ = 1) = 1,$$

where $PPJ = 1$ if we observe a performance payment for the job at least once. So if we observe $PPJ = 1$, we know for sure that it is really a performance-pay job ($PPJ^* = 1$). But when we observe $PPJ = 0$, it may be that the job is or is not one that pays for performance. We have:

$$\begin{aligned} \Pr(PPJ^* &= 1|PPJ = 0, X, T) = \frac{\Pr(PPJ^* = 1, PPJ = 0|X, T)}{\Pr(PPJ = 0|X, T)} \\ &= \frac{q(X) - q(X, T)}{1 - q(X, T)} \end{aligned}$$

which is the fraction of PPJ misclassified as non-PPJ, and

$$\Pr(PPJ^* = 0|PPJ = 0, X, T) = \frac{1 - q(X)}{1 - q(X, T)}$$

Using the wage equations

$$\begin{aligned} W^p &= Xb^p + e^p, \\ W^n &= Xb^n + e^n, \end{aligned}$$

and the assumption that PPJ^* does not depend on the error term (it only depends on X), it follows that:

$$E(W|PPJ = 1, X, T) = E(W|PPJ^* = 1, X) = Xb^p,$$

and

$$\begin{aligned} E(W|PPJ &= 0, X, T) = \Pr(PPJ^* = 1|PPJ = 0, X, T) \cdot E(W|PPJ^* = 1, X) \\ + \Pr(PPJ^* &= 0|PPJ = 0, X, T) \cdot E(W|PPJ^* = 1, X) \\ &= \frac{q(X) - q(X, T)}{1 - q(X, T)} \cdot Xb^p + \frac{1 - q(X)}{1 - q(X, T)} \cdot Xb^n \\ &= Xb^p + \frac{1 - q(X)}{1 - q(X, T)} \cdot X(b^n - b^p) \end{aligned}$$

Finally we can combine this in a pooled model:

$$\begin{aligned} E(W|PPJ, X, T) &= PPJ \cdot Xb^p + (1 - PPJ) \cdot \left(Xb^p + \frac{1 - q(X)}{1 - q(X, T)} \cdot X(b^n - b^p) \right) \\ &= Xb^p + (1 - PPJ) \cdot \frac{1 - q(X)}{1 - q(X, T)} \cdot X(b^n - b^p) \end{aligned}$$

The only difference with a standard pooled model is that the non-performance-pay dummy $(1 - PPJ)$ is being multiplied by an adjustment factor $\frac{1-q(X)}{1-q(X,T)}$. The intuition is very simple. Jobs classified as non-performance-pay are a mix of true non-performance-pay jobs, and performance-pay jobs wrongly classified as non-performance-pay because we have not observed a payment. So instead of moving the coefficient all the way from b^p to b^n when we switch to non-performance-pay, we only go a fraction $\frac{1-q(X)}{1-q(X,T)}$ of the way because some

performance-pay jobs are contaminating the non-performance-pay sample.

In terms of estimation, $q(X,T)$ can be estimated by fitting a logit or linear probability model (with year dummies being part of X). Now, from the equation

$$\Pr(PPJ = 1|X,T) = q(X,T) = q(X)[1-(1-P)^T],$$

it follows that $q(X) = \lim_{T\to\infty} q(X,T)$.

With any reasonable value of P (e.g. $P = .25$, i.e. performance pay received only one year out of four), it follows that with the longest T observed in the data ($T = 22$) we have $q(X, T = 22) \approx q(X)$. So we can run a model (logit or linear probability) to estimate $q(X,T)$ as the predicted probability of performance pay, and get $q(X)$ by replacing the observed value of T with $T = 22$. From this we obtain an estimate of the adjustment factor $\frac{1-q(X)}{1-q(X,T)}$. The average value we obtain is 64 percent, which means that 36 percent of people classified as non-performance-pay would be reclassified as performance pay if we had enough observations on them. So, intuitively, the difference between b^p to b^n should be about 50 percent larger than what we estimate in a pooled model without the adjustment. This is confirmed by comparing the measurement-error adjusted results reported in column 2 of Table B.1 to the benchmark estimates reported in column 1.

Appendix 3: Robustness checks

As discussed in Section IV of the paper, while our measure of performance pay is rather crude, the growth in performance pay is robust to the way we measure performance pay (Table II). We now show that the results for the wage equations (Tables IV) are also robust to these measurement issues.

A first piece of evidence comes from re-estimating all of our models using a number of different specifications and alternative definitions of performance-pay jobs in the PSID. All of our main findings are robust to these specification choices. To summarize this point, we report in Table B.1 OLS and fixed effect estimates of the difference in the returns to education between performance-pay and non-performance-pay jobs.

The benchmark estimates reported in column 1 of Table B.1 correspond to the base estimates reported in columns 3 and 4 of Table IV. Column 2 reports the estimates using 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. The procedure is explained in detail in Appendix 2. Column 3 shows what happens when we allow the return to education to vary across occupations, as in Gibbons, Katz, Lemieux, and Parent (2005). The potential concern here is that the return to education may be higher in performance-pay jobs because they are concentrated in occupations where education is more highly rewarded in the labor market. Column 4 expands on the specifications reported in Table IV by allowing all explanatory variables, as opposed to only the three variables shown in Table IV, to have different effects for the two types of jobs.

Columns 5 to 8 show the estimates using the stricter definitions of performance pay used in Table II, while column 9 adds in workers from the public sector. As we can see from the results reported in columns 1 to 9, both OLS and fixed effect estimates of the return to education are systematically larger for performance-pay than non-performance-pay jobs, and this finding is highly robust across specifications. This is also true if we change our definition of performance-pay jobs and instead simply define performance-pay workers as those having received performance pay in the current year, as is done in column 10. In column 11 we also check whether the results change if we restrict the definition of performance-pay jobs to include only bonuses. While commissions and piece rates are truly individual-based compensation forms, it could be that bonuses are more likely to depend on firm- or group-level performance rather than on individual worker performance. However, we can see that the return to education is also higher in these bonus-pay jobs, and the estimates are very similar to what we find using our benchmark definition.

Finally, in column 12 we estimate the difference in the return to education between performance-pay and non-performance-pay jobs netting out the performance-pay component from total earnings. This gives us a measure of the "base" wage. Interestingly, the higher sensitivity of wages to skills in performance-pay jobs also holds for the base pay component. This suggests that performance pay is a marker for a more flexible compensation package (base wage plus bonus), a result that is consistent with our interpretation that workers are more closely monitored in performance pay jobs, and hence employers are able to adjust total compensation accordingly.

A second piece of evidence in support of our main findings comes from the NLSY data. In the NLSY, workers are directly asked whether or not their earnings are based in part on

their own performance and if so, what form those pay-for-performance payments take (bonus, piece rate, commission). Unfortunately, the question about performance pay in the NLSY was only included in the late 1980s (1988, 1989, and 1990) and late 1990s (1996, 1998, and 2000). Combined with the fact that the NLSY only follows a narrow cohort of individuals over time, it is not possible to use the NLSY to look at the broad effects of performance pay on changes in wage inequality. So we only use the NLSY data to perform some additional robustness checks.

As in the case of the PSID, we focus only on males. We also impose a couple of additional sample restrictions similar to those used by Gibbons, Katz, Lemieux, and Parent (2005). As in the case of the PSID, we classify a job as a performance-pay job when the worker reports performance pay at least once on that job. Note, however, that the limited number of years in which performance pay is measured means that we are less likely to "catch" performance-pay jobs.

As in Table IV for the PSID, we run separate wage regressions for performance-pay and non-performance-pay jobs.[2] 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. Since the AFQT score is purely a measure of worker characteristics, as opposed to job characteristics, its effect on wages should be larger in performance-pay than non-performance-pay jobs. The results, reported in Table B.2, show that both in the late 1980s and late 1990s, returns to productive worker characteristics (education, experience, and the AFQT score) are larger in performance-pay than non-performance-pay jobs, though the differences are not significant in 1988-90. As in the PSID, the return to these characteristics increases over time for both types of jobs, and the increase is, if anything, larger for performance-pay jobs. For example, the return to education increases by 0.0252 for performance-pay jobs (from 0.0703 in column 1 to 0.0955 in column 3), compared to 0.0199 for other jobs (from 0.0550 in column 2 to 0.0749 in column 4).

[2]The wage variable in the estimated models is the hourly wage on the current job at the time of the survey. For simplicity we report regression results with only education, experience, and AFQT as control variables. The results are similar when we use specifications more directly comparable to those used in the PSID.

Appendix 4: Estimation of the variance components model

To see how the variance components model of Section V.3 is identified, consider the expected value of the different cross-products of residuals in the case where the factor loadings (the d's) do not change over time. For individuals on performance-pay jobs, the expected value of the squared residuals is $E(e^p_{ijt} \cdot e^p_{ijt}) = (d^p)^2 \cdot var(\theta_i) + var(\nu^p_{ij}) + var(\varepsilon^p_{ijt})$, the expected value of cross-products for two observations (at time t and time s) on the same job j is $E(e^p_{ijt} \cdot e^p_{ijs}) = (d^p)^2 \cdot var(\theta_i) + var(\nu^p_{ij})$, and the expected value of cross-products for two observations on different jobs j and k is $E(e^p_{ijt} \cdot e^p_{iks}) = (d^p)^2 \cdot var(\theta_i)$. In this simple example, we can estimate the three error components $(d^p)^2 \cdot var(\theta_i)$, $var(\nu^p_{ij})$ and $var(\varepsilon^p_{ijt})$ by taking simple differences of the sample analogs of $E(e^p_{ijt} \cdot e^p_{ijt})$, $E(e^p_{ijt} \cdot e^p_{ijs})$, and $E(e^p_{ijt} \cdot e^p_{iks})$. The same procedure can then be used to estimate the three components $(d^n)^2 \cdot var(\theta_i)$, $var(\nu^n_{ij})$ and $var(\varepsilon^n_{ijt})$ for non-performance-pay jobs. The ratio of the (square) return to unobserved worker characteristics in the two sectors, $(d^p/d^n)^2$, can then be computed as the ratio of the estimated components $(d^p)^2 \cdot var(\theta_i)$ and $(d^n)^2 \cdot var(\theta_i)$.

In the case where factor loading change over time, the expected value of the own-product becomes $E(e^p_{ijt} \cdot e^p_{ijt}) = (d^p_t)^2 \cdot var(\theta_i) + var(\nu^p_{ij}) + \sigma^2_{\varepsilon,t}$, where the factor loading d^p_t and the idiosyncratic variance $\sigma^2_{\varepsilon,t} = var(\varepsilon^p_{ijt})$ are allowed to change over time. The expected value of cross-products for two observations (at time t and time s) on the same job j is $E(e^p_{ijt} \cdot e^p_{ijs}) = d^p_t \cdot d^p_s \cdot var(\theta_i) + var(\nu^p_{ij})$. Since these equations are now non-linear in the parameters (factor loadings), we estimate the models by jointly fitting equations for all the cross-products using non-linear least-squares.

References

Baker, George P., " Incentive Contracts and Performance Measurement," *Journal of Political Economy*, 100 (1992), 598–614.

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.

Bengt Holmström, Bengt, and Paul Milgrom, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55 (1987), 303–328.

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

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

Table B.1
Robustness of Estimates of the Difference in Returns to Education between Performance- and Non-Performance-Pay Jobs

			Additional interactions	
	Benchmark	Measurement error adjusted	Education X occupations	Perf.-pay X all variables
	[1]	[2]	[3]	[4]
OLS	0.0365	0.0439	0.0267	0.0263
	(0.0071)	(0.0100)	(0.0071)	(0.0081)
Fixed effects	0.0169	0.0224	0.0160	0.0107
	(0.0048)	(0.0075)	(0.0051)	(0.0057)
	(Minimum) Frequency of actual performance-pay payments on the job			
	1/5	1/4	1/3	1/2
	[5]	[6]	[7]	[8]
OLS	0.0465	0.0453	0.0498	0.0455
	(0.0079)	(0.0081)	(0.0085)	(0.0095)
Fixed effects	0.0154	0.0153	0.0152	0.0104
	(0.0057)	(0.0057)	(0.0063)	(0.0077)
	Public sector included	PP received this year	Bonus only	Base pay (net of PP)
	[9]	[10]	[11]	[12]
OLS	0.0320	0.0386	0.0379	0.0345
	(0.0065)	(0.0068)	(0.0076)	(0.0075)
Fixed effects	0.0175	0.0059	0.0150	0.0109
	(0.0046)	(0.0040)	(0.0054)	(0.0065)

Notes: Standard errors (in parentheses) are adjusted for clustering at the job-match level. All models are estimated using the specifications used in columns 3 and 4 of Table IV. See the note to Table IV for further details. 26,146 observations used in all models except in column 9 where 32,982 observations are used.

Table B.2
Skills Related Wage Differentials and Performance-Pay (PP) Jobs in the NLSY

	1988-90		1996-2000	
Variable	PP jobs [1]	Non-PP jobs [2]	PP jobs [3]	Non-PP jobs [4]
Years of education	0.0703	0.0550	0.0955	0.0749
	(0.0076)	(0.0044)	(0.0087)	(0.0053)
	[0.082]		[0.042]	
Potential experience	0.0466	0.0440	0.0427	0.0278
	(0.0060)	(0.0031)	(0.0048)	(0.0029)
	[0.704]		[0.008]	
AFQT score (/10)	0.0416	0.0330	0.0534	0.0443
	(0.0057)	(0.0032)	(0.0071)	(0.0039)
	[0.187]		[0.0266]	
Number of observations	1,553	4,726	1,053	2,870

Notes: Regression models estimated for males only. Robust standard errors in parentheses.
P-values of tests of equality of coefficients (PP jobs vs. other jobs) in square brackets.

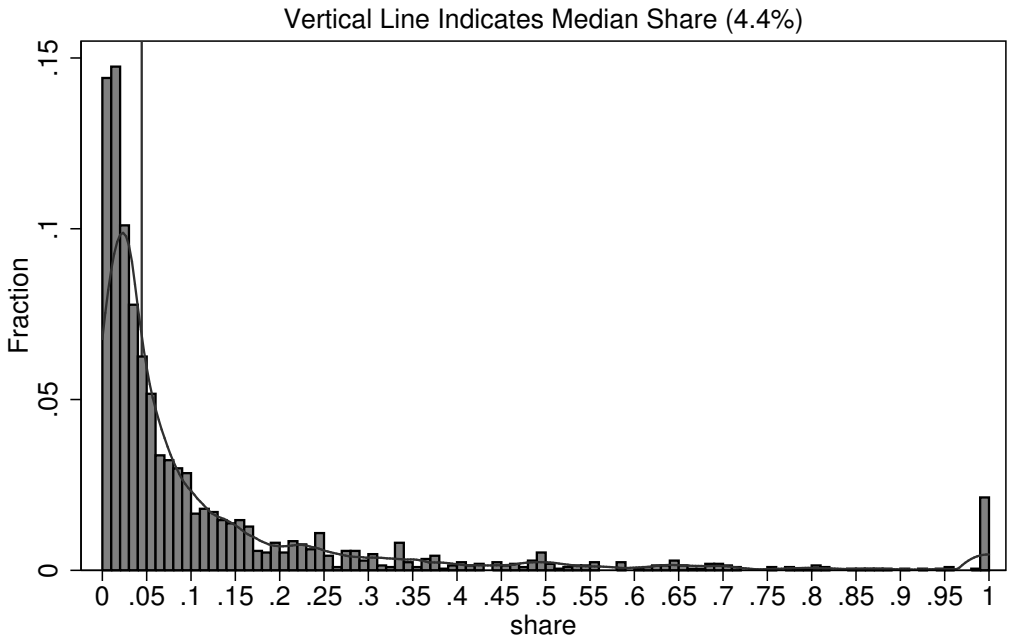

Figure B.1
Share of Performance Pay in Total Earnings:PSID 1976−1998