

NBER WORKING PAPER SERIES

MINIMUM WAGES AND  
ON-THE-JOB TRAINING

Daron Acemoglu  
Jörn-Steffen Pischke

Working Paper 7184  
<http://www.nber.org/papers/w7184>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
June 1999

We thank John Browning and especially Aimee Chin for excellent research assistance, and Joe Altonji, Richard Carson, Ken Chay, Jinyong Hahn, Lisa Lynch, and Chris Taber for useful comments. The first draft of this paper was written while Pischke was a visiting scholar at the Northwestern University/University of Chicago Joint Center for Poverty Research. He is grateful for their hospitality and research support. Acemoglu is grateful for financial support under the National Science Foundation Grant SBR-9602116. All opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1999 by Daron Acemoglu and Jörn-Steffen Pischke. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Minimum Wages and On-the-job Training  
Daron Acemoglu and Jörn-Steffen Pischke  
NBER Working Paper No. 7184  
June 1999  
JEL No. J24, J31, J41

**ABSTRACT**

Becker's theory of human capital predicts that minimum wages should reduce training investments for affected workers, because they prevent these workers from taking wage cuts necessary to finance training. We show that when the assumption of perfectly competitive labor markets underlying this theory is relaxed, minimum wages can increase training of affected workers, by inducing firms to train their unskilled employees. More generally, a minimum wage increases training for constrained workers, while reducing it for those taking wage cuts to finance their training. We provide new estimates on the impact of the state and federal increases in the minimum wage between 1987 and 1992 of the training of low wage workers. We find no evidence that minimum wages reduce training. These results are consistent with our model, but difficult to reconcile with the standard theory of human capital.

Daron Acemoglu  
Department of Economics  
Massachusetts Institute of Technology  
Cambridge, MA 02139  
and NBER  
daron@mit.edu

Jörn-Steffen Pischke  
Department of Economics  
Massachusetts Institute of Technology  
Cambridge, MA 02139  
and NBER  
jsp@mit.edu

# 1 Introduction

Much of the recent debate on the minimum wage has focused on its employment implications. The theory of human capital suggests that minimum wages should also have important adverse effects by reducing training and wage growth. In the standard human capital theory, as developed by Becker (1964), Ben-Porath (1967), and Mincer (1972), a large part of human capital is accumulated on the job, and workers often finance these investments through lower wages. A binding minimum wage will therefore reduce workplace training, as it prevents low wage workers from accepting the necessary wage cuts (Rosen, 1972). The early empirical literature has confirmed this prediction. The negative impact on human capital formation has been an important argument against minimum wages in the minds of many economists and policy-makers.

In this paper, we revisit the impact of minimum wages on training. We provide new empirical estimates that are quite different from those already in the literature, bringing a new perspective to this debate. We also provide some new theoretical ideas on the impact of minimum wages on training, which imply that minimum wages may increase the training of low paid workers. Since the standard theory of human capital predicts a negative effect of minimum wages on training, our evidence enables an evaluation of the empirical success of the alternative theoretical approaches.

The standard theory of human capital is based on competitive labor markets. The implications of minimum wages in noncompetitive labor markets are very different. We show that in the presence of labor market frictions, minimum wages shift the incentives for training investments from the worker to the employer, and may increase training investments. The intuition for this result is that minimum wages make it less profitable to employ unskilled workers. When there are no rents in the employment relationship, as in a competitive labor market, the firm has no option but to lay off workers who were previously paid below the new minimum wage. In contrast, in the presence of labor market rents, it may be more profitable to increase the productivity of workers, who are already receiving high wages, rather than laying them off. Figure 1 illustrates this intuition diagrammatically. It draws the relation between worker skills,  $\tau$ , productivity,  $f(\tau)$ , and wages  $w(\tau)$ . The gap between productivity and wages,  $\Delta$ , is the rent that the firm obtains. A binding minimum wage, in the absence of such rents, forces the firm to lay off the worker. However, with  $\Delta$  sufficiently high, the firm would like to retain the worker despite the higher wages dictated by the minimum wage. In this case, the firm would also like to increase the productivity of the worker. Without the minimum wage, the gap between  $f(\tau)$  and  $w(\tau)$  was constant, so there was no point in incurring costs of training. However, with a minimum wage, profits are less at  $\tau = 0$  than at  $\tau = 1$ . So

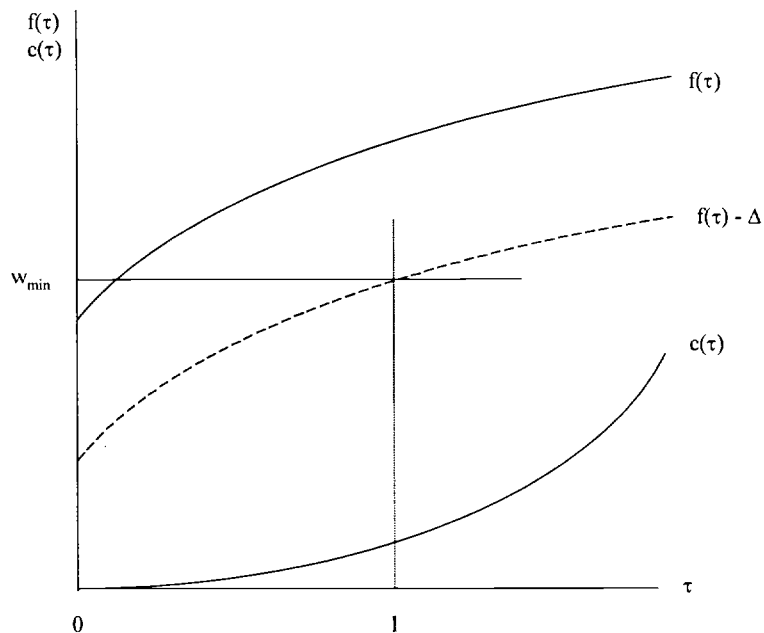


Figure 1: Training with a Minimum Wage and Employment Rents

if the firm can increase its employee's skills to  $\tau = 1$  at a moderate cost, it will prefer to do so. In essence, the minimum wage has made the firm the *de facto* residual claimant of the increase in the worker's productivity, whereas without the minimum wage, the worker was the residual claimant.

This reasoning suggests that a binding minimum wage may induce firms to invest more in the skills of their employees. As in the standard theory, however, the wages of some workers may have been low because they were compensating their employers for investments in general skills, and the minimum wage would prevent this. Our approach, therefore, predicts that minimum wages reduce the training of workers taking wage cuts to finance their training, while inducing further training for those who were constrained in their human capital investments. In the first part of the paper, we outline this theory in greater detail.

The broad aggregate trends in the data do not appear in line with the standard theory. A variety of evidence, including the spike in the wage distribution at the minimum (e.g. DiNardo, Fortin, and Lemieux, 1996; Card and Krueger, 1995), suggests that minimum wages are binding for some workers. Therefore, the standard theory predicts that the decline in the real minimum wage during the 1980s should have increased training for low wage workers. Comparing formal and informal company training among high

school dropouts in the Training Supplements of the CPS, we find this not to be the case. The training incidence falls from 15 percent in 1983 to 11 percent in 1991, even though the real value in the minimum wage was lower in 1991 than in 1983 (see Acemoglu and Pischke, 1999a).

Nevertheless, other aggregate trends may have been responsible for the decline in training among low wage workers. In our micro data work, we control for these aggregate trends and exploit cross-state and region variations in how binding minimum wages are. We use the National Longitudinal Survey of Youth (NLSY) for the period 1987 to 1992. This period encompasses a number of state minimum wage increases as well as two federal increases in the minimum in 1990 and 1991. Our data therefore contain a large amount of within state variation in minimum wages. Furthermore, the NLSY is a panel of youths and oversamples those from disadvantaged backgrounds, so it contains a relatively high number of low wage workers directly affected by minimum wage increases.

Our empirical results show almost no evidence of a reduction in training in response to minimum wages. Some of our specifications, which look most directly at workers affected by minimum wage increases, suggest that minimum wages may have increased worker training slightly, although these positive coefficients are not statistically significant.

No (or small positive) effects of minimum wages on training investments for workers for whom the minimum is binding suggest that the adverse effects of minimum wages on the human capital of low wage workers may be rather limited. These findings contrast with the predictions of the standard theory. The standard theory implies that an increase in the minimum wage should eliminate all general training among affected workers, as these investments have to be financed solely by the workers themselves. Our results are therefore even harder to reconcile with the standard theory if most training in the NLSY is general, as Loewenstein and Spletzer (1997) find to be the case in later waves. Our evidence therefore suggests that models with imperfect labor markets may be more appropriate for an analysis of training decisions and the impact of minimum wages on training investments than models based on competitive labor markets.

The rest of the paper is organized as follows. The following section discusses the related literature. Section 3 presents a simple theoretical setup where, contrary to the predictions of the standard theory, minimum wages increase training investments. This setup is special and is chosen to illustrate the main theoretical contribution of this paper clearly. In Section 4, we consider a more general and realistic model where minimum wages increase training for some workers, while reducing it for others. In Section 5, we describe our empirical strategy and give an overview of the data. Our results are presented in Section 6 and Section 7 concludes.

## 2 Previous Literature

Previous theoretical discussions of the effect of the minimum wage on training are based on the human capital model of Becker (1964). In the models of Rosen (1972) and Hashimoto (1982), minimum wages reduce training investments, because they prevent workers from taking the wage cuts necessary to finance general human capital investments. Our theoretical result deviates sharply from this conventional wisdom. We build on our previous work, Acemoglu and Pischke (1999b), which showed that a compression in the structure of wages can induce firm-sponsored training. Here, we show that minimum wages shift training investments from workers to firms, and may increase the overall amount of training.

There is a small empirical literature investigating the impact of minimum wages on training. Part of this literature focuses on whether minimum wage laws lead to slower observed wage growth in micro data. Both Leighton and Mincer (1981) and Hashimoto (1982) have found this to be the case and concluded that minimum wage laws lead to less training. In our model, the introduction of a small minimum wage could lead to a shift of training expenses from workers to firms, and reduce wage growth, even when it causes an increase in training. More generally, since minimum wages cut the lower tail of the wage distribution, and typically create a spike at the minimum, they would appear to reduce age-wage profiles even when they have no effect on training. Therefore, it is unsatisfactory to interpret the decline in age-earnings profiles as evidence of reduced investment in general training. Consistent with this view, Grossberg and Sicilian (1999) find no effect of minimum wages on training, but still find lower wage growth for minimum wage workers. Furthermore, Card and Krueger (1995) compared cross sectional wage profiles in California before and after the 1988 minimum wage increase with a number of comparison states. They also found flatter profiles in California after the minimum wage increase. However, they point out that the Californian profile also shifts up and does not cross the previous age-wage profile. This pattern contradicts the standard theory, but is consistent with the predictions of our model.

Given the difficulty of interpreting changes in the slope of wage profiles, we find it more compelling to look at the impact of minimum wages on training directly, but we are only aware of four previous studies doing this for the US. Leighton and Mincer (1981) use worker reported data on the receipt of training from the Panel Study of Income Dynamic (PSID) and the NLS and find that workers in states with lower wages and therefore a more binding federal minimum wage receive significantly less training. Cross state comparisons may be confounded by the presence of other state effects, however. For example, industrial and occupational composition of employment

varies substantially across states, and different industries and occupations have different skill requirements. These considerations suggest that across state comparisons are hard to interpret. Schiller (1994) reports a similar finding using later data from the NLSY by comparing the training incidence of minimum wage workers with those earning higher wages. The evidence from this study is even harder to interpret because worker traits which lead to higher pay are typically also associated with more training. Grossberg and Sicilian (1999) use data from the Employment Opportunity Pilot Project (EOPP) and compare minimum wage workers both to workers earning slightly less and slightly more, ameliorating the problem of worker heterogeneity somewhat. They find insignificant negative effects on training for male minimum wage workers and insignificant positive effects for women. Leighton and Mincer only analyzed men, although women make up the majority of minimum wage workers.

Some of these problems are overcome in a more recent study by Neumark and Wascher (1998), who use Current Population Survey (CPS) supplements to compare the impact of minimum wages on training within states using comparisons of young workers in 1991 with older workers (who are less likely to be affected by the minimum wage) and with young workers in 1983. These comparisons assume that state differences in training levels are the same for younger and older workers and remain so over long time periods, which are stringent requirements. They also find negative effects of minimum wages on training, but these effects seem to be too large to be sensible. Their treatment group consists of all young workers, but not all of these workers are affected by the minimum wage. Their estimates imply that formal training among workers aged 20-24 in California (a high minimum wage state) was 3.2 percentage points lower than in states which were subject to the federal minimum. Suppose this reduction in training is entirely due to workers being affected by the minimum wage. Assuming, quite generously, that minimum wages affect all workers earning less than 160 percent of the minimum (which comprises 30% of workers aged 20-24), this estimate implies that among affected workers, training is lower by over 10 percentage points (i.e. 3.2 percentage points divided by 0.30). It is important to note that the average incidence of training among workers aged 20-24 earning 160 percent of the minimum or less in low minimum wage states is 2.7 percent, much lower than among all workers aged 20-24 (for whom the incidence is 10 percent). So this estimate implies that introducing California's minimum wage to low minimum wage states should have wiped out all training *four times* among affected workers in these states. This suggests that the fixed effects assumptions made by Neumark and Washer are quite suspect.

### 3 Minimum Wages and Training: Basic Theory

In this section, we use a simple two-period model to analyze the impact of minimum wages on training. The main result of this analysis is that plausible deviations from perfectly competitive labor markets, which introduce firm-specific rents, change the conclusions of Becker's seminal analysis. Namely, we find that with labor market imperfections, minimum wages can increase investments in general training, and in most cases, they will increase training for employees previously constrained in their human capital investments, while reducing it for those taking wage cuts to finance their training.

We start with the simplest model which illustrates our most original result. In this model, all skills are general, and workers cannot finance their own training. With competitive labor markets and without minimum wages, there are no training investments. As a result, moderate minimum wages increase training unambiguously.

These simplifying assumptions are unnecessary for our main insights, however, and they hide the potential adverse effects of minimum wages on training. In the next section, we consider a less restrictive model where workers are potentially heterogeneous, and can take wage cuts to finance training. In this more general environment, minimum wages will increase training for some workers, but reduce it for those previously taking wage cuts to finance training.

#### 3.1 Environment

The world lasts for two periods, 1 and 2. There is no discounting, and all agents are risk-neutral. There is a continuum of workers with density  $n$ , who supply labor inelastically. We start with the case in which all workers have common ability normalized to 1. We view period 1 as the early career of workers. During this period a worker's productivity is  $\theta < 1$ , and he can also receive training. To simplify the discussion we assume that training is indivisible, so only  $\tau = 0$  (no training) and  $\tau = 1$  (training) are possible. Without training, but with the experience of first period employment, the worker's productivity increases to 1 in period 2. Training, in turn, increases his productivity in period 2 further, to  $\phi > 1$ . We assume that training is general, so this increase in productivity applies equally in all firms. The cost of training is incurred in terms of lower output in the first period and is equal to  $c > 0$ . As noted above, in this section, we assume that all training investments have to be financed by the firm. This could be because training is non-contractible, so the firm can renege on its training promise even if the worker takes a wage cut to finance his training (see Acemoglu and Pischke, 1999a, for a discussion).

We now impose:



**Assumption 1:**  $\phi - 1 > c$ .

This condition ensures that training is socially beneficial.

There is a continuum 1 of firms competing to hire workers in both periods. Each firm can hire at most one worker. Firm  $j$  has to incur a setup cost  $k_j$  (e.g. buy some equipment). We assume that  $k_j$  has a continuous distribution among firms given by  $F(k)$  with lower support at  $\underline{k} < 0$  and upper support at  $\bar{k} < \delta + \theta$  ( $\delta$  is defined below). This smooth distribution enables us to generate a downward sloping demand for labor in period 1. A negative value of  $k$  implies that the firm will make profits in the first period even after incurring the required capital cost, for example because it has some special expertise in this area of business. However, to realize these profits, it needs to hire a worker and to produce. We suppose that the lower support  $\underline{k}$  is low enough (e.g.  $\underline{k} \rightarrow -\infty$ ), so that minimum wages never stop all employment in the first period.<sup>1</sup>

Once set up, all firms have access to the common technology. In the first period, revenue per worker is  $\theta$ . Their second period revenue is 1 if they employ an untrained worker and  $\phi$  if they employ a trained worker. Also, the firm obtains an additional revenue  $\delta$ , if the worker was employed in the first period with this firm. Therefore,  $\delta$  is a firm-specific productivity increase, and is the only deviation from competitive labor markets we introduce.  $\delta$  creates a match-specific surplus to be divided between the firm and the worker. We assume Bertrand competition among firms (or equivalently, firms make take it or leave it offers to workers). Since the second-best opportunity of workers does not include  $\delta$ , initial employers capture the whole amount  $\delta$ , only paying the “market wage”, even though workers’ productivity is higher than this. As will become apparent shortly, the presence of rents from the employment relationship is crucial, but the results would be unaffected if the worker captures a fraction  $\beta < 1$  of this surplus. Observe also that this mobility cost implies that all workers will stay with their initial employer.<sup>2</sup>

---

<sup>1</sup>Decreasing returns at the firm level is an alternative assumption which generates a smoothly decreasing demand for labor. The equivalent of the assumption  $\underline{k} \rightarrow -\infty$  in that case is that the marginal product of labor in the first period is very large when very few workers are employed. This formulation yields the same results.

<sup>2</sup>Alternatively, we could have that a worker who changes a job in the second period incurs a mobility cost  $\delta$ , because he has to acquire some costly firm-specific skills before becoming productive. All our results continue to hold with this alternative interpretation. Also, in practice, workers earning close to the minimum wage have high mobility. Our results continue to hold if a fraction  $s < 1$  of workers change jobs.

### 3.2 Equilibrium Without Minimum Wages

Equilibrium can now be characterized by backward induction. In the second period, each firm is willing to pay up to 1 for an untrained worker employed by another firm, and is willing to pay  $\phi$  for a trained worker. Bertrand competition then ensures that<sup>3</sup>

$$\begin{aligned} w_2(NT) &= 1, \\ w_2(T) &= \phi. \end{aligned} \tag{1}$$

Using the fact that in equilibrium all workers stay with their first period employer, the profits that firm  $j$  makes from an untrained and trained worker can be written as:

$$\begin{aligned} \Pi_j(NT) &= (\theta - k_j - w_1) + (1 + \delta - w_2(NT)) \\ \Pi_j(T) &= (\theta - k_j - c - w_1) + (\phi + \delta - w_2(T)) \end{aligned} \tag{2}$$

The first bracket in each expression gives first period profits, while the second is profits in period 2. Substituting (1) into (2), we immediately see that  $\Pi(T) - \Pi(NT) = -c$ , that is the firm loses the training costs if it trains. The reason is simple: a trained worker receives  $\phi - 1$  more in the second period, and this is exactly the increase in his contribution to the firm's output. So the worker is the full residual claimant of the increase in productivity due to training, and the second period profit of the firm is equal to  $\delta$ , independent of whether the worker is trained. Since training costs  $c$  in the first period has to be paid by the firm, it is not optimal for firm to invest in its employee's skills. Although there are firm-specific rents, because they do not interact with training, firms have no incentive to train (see Acemoglu and Pischke, 1999b). Therefore, with competitive labor markets, firms have no incentive to invest in their workers' general skills, even though training is socially desirable.

To complete the characterization of the equilibrium, we have to determine first period wages. It is clear that only the  $n$  firms with the lowest capital costs will be active. So, the firm with the  $n^{\text{th}}$  lowest capital cost  $k_n$  has to make zero profits. This gives:

$$w_1 = \theta + \delta - k_n \tag{3}$$

A noteworthy feature is that compensation can be front-loaded or back-loaded. Front-loading arises because firms anticipate  $\delta$ , the rent they will receive in the second period, and are willing to bid higher than the worker's current productivity. The reason why compensation could be back-loaded is that firms incur the capital cost only in the first

---

<sup>3</sup>Even though workers do not receive any of the firm-specific rents, there is no labor mobility in equilibrium. Suppose this were not so, then the firm could offer  $\varepsilon$  more than the market wage and convince the worker to stay, so mobility cannot be part of an equilibrium.

period, and anticipate to compete harder for workers in the second period. We assume that

**Assumption 2:**  $\theta + \delta - k_n < 1$ .

This limits front-loading and ensures that there is a positive experience premium even without training, i.e.  $w_1 < w_2(NT)$ . Recall that  $k_n$  is not an endogenous variable, but a parameter—the  $n^{\text{th}}$  lowest capital cost.

Now, denoting total employment by  $E$ , we can summarize this analysis (proof in the text):

**Proposition 1** *There is a unique equilibrium in which there is no training and all firms with  $k_j \leq k_n$  are active. All workers are employed,  $E = n$ , and receive the first period wage given by (3) and the second period wage  $w_2(NT)$  given by (1).*

So despite Assumption 1, which ensures that training is socially beneficial, there is no training in equilibrium. This is in line with Becker’s standard theory; firms are unwilling to invest in the general skill of their employees as they do not receive any of the benefits of training, and workers are assumed unable to “buy” training. However, there is full employment in this decentralized economy, in particular, the level of employment,  $E$ , is equal to the labor force,  $n$ .

### 3.3 The Impact of Minimum Wages on Training

Now consider the imposition of a binding minimum wage  $w_M > 1$ . Assumption 2 ensures that  $w_M > \theta + \delta - k_n$ , so the minimum wage is also binding in the first period. Since the demand for labor is downward sloping, the minimum wage will reduce employment in this economy. The main question for the focus of this paper, however, is how the minimum wage affects training.

We start by writing wages in the presence of minimum wage laws. With a similar reasoning to above, (1) and (3) change to:

$$\begin{aligned} w_1 &= w_M, \\ w_2(NT) &= w_M, \\ w_2(T) &= \max\{w_M; \phi\}. \end{aligned} \tag{4}$$

$w_2(T)$  features the “max” operator because the minimum wage may be less than or greater than the productivity of a trained worker,  $\phi$ . Profits with and without training

now become:

$$\begin{aligned}\Pi_j(NT) &= (\theta - k_j - w_M) + \max[(1 + \delta - w_M); 0] \\ \Pi_j(T) &= (\theta - k_j - c - w_M) + \max[(\phi + \delta - \max\{w_M; \phi\}); 0]\end{aligned}$$

where the first “max” operator takes care of the fact that the firm may decide not to operate in the second period if the minimum wage is too high.

First, suppose  $w_M > \phi$ , so that the second period wage is equal to the minimum even for skilled workers. In this case, we have

$$\Pi_j(T) - \Pi_j(NT) = \phi - 1 - c > 0,$$

which is always positive by Assumption 1. So in stark contrast to the economy without minimum wages, firms now prefer to train their employees. The firm has to pay the minimum wage irrespective of whether the worker is skilled or not, so the full return to training is captured by the employer.

Next, suppose that  $w_M < \phi$ , so that second period wage for trained workers exceeds the minimum. Then,  $\Pi_j(T) - \Pi_j(NT) = w_M - 1 - c$ . This can be positive or negative depending on how high the minimum wage is.

At this point, notice the importance of firm-specific rents. If  $\delta = 0$ , then in the second period, a firm which trains pays  $\max\{\phi, w_M\}$ , so necessarily loses money. So there cannot be any training. More generally, in all this analysis we have to make sure that minimum wages are not high enough to shut down all employment in the second period. This requires  $w_M < \phi + \delta - c$ .

We can now state the main result of this section (proof in the text):<sup>4</sup>

**Proposition 2 .**

1. *If  $\phi + \delta - c > w_M > \phi$ , then all employed workers are trained,  $w_1 = w_2(NT) = w_2(T) = w_M$ , and all firms with  $k \leq k^*$  are active where  $k^* \equiv \theta + \phi + \delta - c - 2w_M < k_n$ . Employment in both periods is equal to  $E = 1 - F(k^*)$ .*
2. *If  $1 + c < w_M < \phi$ , then all employed workers are trained,  $w_1 = w_2(NT) = w_M$ , and  $w_2(T) = \phi$ . All firms with  $k \leq k^{**}$  are active where  $k^{**} \equiv \theta + \delta - c - w_M \in (k^*, k_n)$ , so employment in both periods is now  $E = 1 - F(k^{**})$ .*
3. *If  $w_M < 1 + c$ , then there is no training,  $w_1 = w_2(NT) = w_M$ , and  $w_2(T) = \phi$ , all firms with  $k \leq k^+$  are active where  $k^+ \equiv 1 + \theta + \delta - 2w_M \in (k^{**}, k_n)$ , so employment is now equal to  $E = 1 - F(k^+)$ .*

---

<sup>4</sup>When there is training,  $w_2(NT)$  gives “off-the-equilibrium” path wages, and similarly for  $w_2(T)$ , when there is no training. Also notice that if  $c > \delta$ , then we will never be in case 2.

As in the standard neoclassical model, minimum wages reduce employment in all cases, because some firms (jobs) making close to zero profits before the introduction of the minimum will choose to become inactive. But moderate minimum wages increase training. Firm-specific rents (labor market imperfections) are crucial for this result. If  $\delta = 0$ , there is no training and employment in the second period collapses to zero. The firm has to pay minimum wages even for unskilled workers and a binding minimum wage reduces the rents the firm receives from the employment relationship. Training raises the worker's productivity, and therefore restores some of these rents.

Figure 1 in the introduction gives the basic intuition. The minimum wage determines the wage both at  $\tau = 0$  and  $\tau = 1$ , so all productivity increases from training accrue to the employer, as in case 1. The most interesting case might be case 2. If the minimum wage is low enough, it may induce the firm to sponsor training but the worker receives some of the proceeds from the training, because the productivity of the trained worker exceeds the minimum wage.<sup>5</sup>

Notice how minimum wages affect wage growth here. In case 1, the minimum wage induces training, but will lead to a reduction in wage growth (from  $1 - \theta - \delta + k_n$  to zero), so inferring whether minimum wages increase training by looking at the slope of the wage profile, as in Hashimoto (1982), is not possible. Since minimum wages increase first period wages, they will typically reduce wage growth (though in case 2, if  $\phi$  is sufficiently large, they may increase wage growth).

Finally, notice that the results continue to apply if  $w_M$  is a wage above the legal minimum that the firm has to pay to the worker, due to other imperfections such as bargaining. Therefore, if an increase in the minimum creates spillover effects to wages above the minimum, our analysis predicts that firms may also be induced to train the workers affected by these spillovers. In practice, minimum wages appear to create spillover effects (e.g. DiNardo, Fortin and Lemieux, 1996; Lee, 1999), so we expect them to also influence the training of low wage workers earning above the minimum.

## 4 The Effect of Minimum Wages When Workers Can Pay For Training

We now generalize the model of the previous section in two crucial respects: (1) workers can invest in their general training, but some of them are affected by credit constraints; (2) returns to training vary across workers.

---

<sup>5</sup>This is special to the case where the firm has to choose a discrete level of training. If the firm could choose training continuously, it would never train the worker beyond the point where it has to pay above the minimum wage in the second period (see Figure 1).

More specifically, there is a distribution of abilities across workers denoted by  $G(\eta)$  with support  $[\underline{\eta}, \bar{\eta}]$ . A worker with ability  $\eta_i$  produces  $\theta\eta_i$  in the first period,  $\eta_i$  in the second period if untrained, and  $\phi\eta_i$  if trained. The cost of training is  $c$  independent of ability.

To simplify the discussion, we also assume:

**Assumption 3:**  $(\phi - 1)\underline{\eta} > c$ .

This implies that training is beneficial for all workers, so we do not have to keep track of whether the worker in question should get training or not.

Furthermore, we assume that workers can “buy” training from their employers, so the contractual problems discussed in the previous section are absent. Nevertheless, workers do not have access to perfect credit markets, and need to achieve a certain level of consumption in the first period. We denote this by  $\gamma_i$  for worker  $i$ , and denote the distribution of  $\gamma$  by  $H(\gamma)$ , with upper support  $\bar{\gamma}$ . The distribution of  $\gamma$  is independent of the distribution of  $\eta$ . A worker needs to have a consumption level  $\gamma$  out of his wage income, and  $\gamma < 0$  implies that the individual has other assets, so he does not need his wage income to meet his consumption obligations.<sup>6</sup> The rest of the setup is unchanged.

## 4.1 Equilibrium Without Minimum Wages

Decentralized equilibrium without minimum wages is straightforward to characterize. Once again all firms with  $k_j \leq k_n$  are active. Market wages are now conditioned on ability as well as training:

$$\begin{aligned} w_2(\eta_i, NT) &= \eta_i \\ w_2(\eta_i, T) &= \phi\eta_i. \end{aligned} \tag{5}$$

Writing profits with and without firm-sponsored training as before, we see once again that  $\Pi_j(T) - \Pi_j(NT) = -c$ . Firms therefore have no incentive to invest in their employees’ skills. But now, there can be worker-financed training. More specifically, by Assumption 3, all workers who can afford it will prefer to get training by taking a wage cut. Therefore, worker  $i$  gets training if and only if

$$\theta\eta_i + \delta - k_n \geq c + \gamma_i. \tag{6}$$

This inequality ensures that the first period wage is high enough that the worker can afford to pay for his own training. Summarizing this result (proof in the text):

---

<sup>6</sup>This amounts to lexicographic preferences, whereby the worker first tries to meet his minimum consumption level, and then maximizes the sum of consumption in both periods. Introducing a concave utility function for the worker in both periods gives similar results, as it creates a desire for smooth consumption, but complicates the analysis substantially, so we have opted for this simpler formulation.

**Proposition 3** *There is a unique equilibrium in which all firms with  $k_j \leq k_n$  are active, wages in the second period are given by (5), and all workers for whom (6) does not hold receive the first-period wage  $w_1(\eta_i) = \theta\eta_i + \delta - k_n$  and do not obtain training. Workers for whom (6) holds receive the first-period wage  $w_1(\eta_i) = \theta\eta_i + \delta - k_n - c$  and obtain training.*

So in this competitive equilibrium there is investment in training. All training is of the Becker-type, financed by workers taking a wage cut in the first period. The only reason why the equilibrium is not first-best is because some of the workers are credit constrained and cannot invest in training even though doing so would increase total output.

It is also possible to obtain a relatively simple expression for the incidence of training. Because of Assumption 3, all workers who can afford it will invest in training. We can write the incidence of training as

$$T = \int_{\underline{\eta}}^{\bar{\eta}} H(\theta\eta + \delta - k_n - c) dG(\eta).$$

Only workers for whom (6) holds obtain training, and this depends on the size of  $\gamma$  relative to first period wages which depend on their ability  $\eta$ .

## 4.2 The Impact of Minimum Wages on Training

Now consider a binding minimum wage, i.e.  $w_M$  such that  $w_M > \theta\underline{\eta} + \delta - k_n$  and  $w_M > \underline{\eta}$ . Also to reduce the cases to be studied, suppose that  $w_M < \phi\underline{\eta} + \delta - c$ , so that second period employment does not collapse. Similar reasoning to above immediately gives second period wages as:

$$\begin{aligned} w_2(\eta_i, NT) &= \max\{w_M, \eta_i\} \\ w_2(\eta_i, T) &= \max\{w_M, \phi\eta_i\} \end{aligned} \tag{7}$$

Let us start with the adverse effect of the minimum wage on investment in skills. Some of the workers previously financing their own training are now unable to do so. For example, consider worker  $i$  such that  $\gamma_i < \theta\eta_i + \delta - k_n - c < w_M$ , that is worker  $i$  was financing his training by taking a wage cut, to a level now below the minimum. Clearly, the minimum wage prevents this.<sup>7</sup> So a binding minimum wage reduces training for some workers. If  $\delta = 0$ —so that we are in competitive labor markets—, this is the only

---

<sup>7</sup>This is without a training subminimum. In our model, as in the standard Becker model, introducing training subminima increases training.

impact of the minimum wage on training; since all training is general in this economy, the minimum wage eliminates all training among affected workers.

In addition, however, the minimum wage induces firms to train their low skill workers as in the previous section. Consider a worker with  $\gamma_i > \theta\eta_i + \delta - k_n - c$ , that is a worker who could not afford to buy training before the minimum wage. It is then clear, from the analysis in the previous subsection, that if  $w_M > \phi\eta_i$ , then the firm would prefer to invest in training. Similarly, even when  $w_M < \phi\eta_i$ , if  $w_M > \eta_i + c$ , the firm prefers to provide training. Therefore, as in the previous section, labor market rents and minimum wages induce employers to invest in general training of some workers.

More formally, as in the previous section, a worker receives firm-sponsored training if

$$w_M > \eta_i + c \tag{8}$$

(since  $\phi\eta_i > \eta_i + c$ , if the trained worker's wage is at the minimum, (8) is still satisfied). Alternatively, the worker will obtain training when he finances it himself, despite the minimum wage. This happens if

$$\theta\eta_i + \delta - k^* - c > \max\{w_M, \gamma_i\} \tag{9}$$

where  $k^*$  is the new equilibrium cutoff capital cost.

These observations provide some insight into the effects of minimum wages on training. Minimum wages induce firms to train their workers, but prevent workers' own investment in training. To understand the whole picture, however, we need to characterize the equilibrium fully, which is a little more involved than before. We have to determine first period wages for workers of different abilities, and find the equilibrium level of employment when a minimum wage is binding. However, since first period wages are not central to our focus, we only give these in Appendix 1, and only determine equilibrium employment here.

The key observation is that there exists, once again, a cutoff level of capital cost, such that all  $k_j \leq k^*$  will be active. Furthermore, given the additive relation between firm and worker heterogeneity, firm  $j$ , if active, has to make exactly the same total profit,  $\Pi_j$ , from hiring any worker  $i$ , so that each firm is indifferent about which worker to hire (otherwise, all firms would want to hire the more profitable worker, pushing up his period 1 wage). So let us write the profit that a firm makes from a worker of talent  $\eta_i$  as:

$$\Pi_j(\eta_i) = \max \{ \Pi_j(\eta_i, NT), \Pi_j(\eta_i, T) \},$$

where

$$\Pi_j(\eta_i, NT) = (\theta\eta_i - w_1(\eta_i, NT) - k_j) + (\eta_i + \delta - \max\{w_M, \eta_i\}),$$



$$\Pi_j(\eta_i, T) = (\theta\eta_i - w_1(\eta_i, T) - k_j - c) + (\phi\eta_i + \delta - \max\{w_M, \phi\eta_i\})$$

after substituting (7).  $w_1(\eta_i, NT)$  denotes the first period wage for a worker with ability  $\eta_i$  not receiving training, and  $w_1(\eta_i, T)$  denotes the first period wage for a worker receiving training. These first period wages have to adjust to ensure that  $\Pi_j(\eta_i) = \Pi_j = k^* - k_j$  for all  $i$  hired in equilibrium, where  $k^*$  is the equilibrium cutoff capital cost. This expression follows by noting that the firm with  $k^*$  will make zero profits, so a firm with  $k_j < k^*$  should make profits equal to  $k^* - k_j$ . Clearly  $w_1(\eta_i, \cdot) \geq w_M$  by law.

To characterize the equilibrium, we need to determine  $k^*$ . Since the minimum wage is binding, the marginal worker has to be employed at the minimum.<sup>8</sup> Hence, defining the ability of the marginal worker as  $\eta^*$ , we must have

$$\Pi_{j^*}(\eta^*) = (\theta\eta^* - k^* - w_M - c) + (\phi\eta^* + \delta - \max\{w_M, \phi\eta^*\}) = 0,$$

where  $k_{j^*} = k^*$ , and we have used the fact that a worker at the minimum wage in the second period necessarily receives training. Hence:

$$k^* = (\theta + \phi)\eta^* + \delta - c - w_M - \max\{w_M, \phi\eta^*\}. \quad (10)$$

The number of workers employed, i.e. those with  $\eta \geq \eta^*$ , has to be equal to the number of active firms, i.e. to the firms with  $k \leq k^*$ . Hence:

$$[1 - G(\eta^*)]n = F(k^*) \quad (11)$$

where  $k^*$  is as given by (10). Notice that the left hand side is strictly decreasing in  $\eta^*$ , while the right hand side is increasing. Together with the boundary conditions implied by the fact that  $G$  and  $F$  are distribution functions, this ensures a unique solution  $(\eta^*, k^*)$ .

We can now summarize the results (proof in the text):

**Proposition 4** *There is a unique equilibrium with employment level  $E = F(k^*)$  where  $k^*$  is given by (10) and (11). Workers with  $\eta_i$  satisfying (8) or (9) receive training.*

The incidence of training among employed workers is now:

$$T^{MW} = \frac{n}{F(k^*)} \left[ G(w_M - c) - G(\eta^*) + \int_{\underline{\gamma}}^{\bar{\gamma}} [1 - G(\hat{\eta}(\gamma))] dH(\gamma) \right]$$

---

<sup>8</sup>Alternatively, the marginal worker could be one who has lower talent but low  $\gamma$ , so that he can pay for his training. This would imply  $\eta_i \geq \tilde{\eta} = [w_M + c + k^* - \delta] / \theta$ . However,  $\tilde{\eta}$  is unambiguously greater than  $\eta^*$  obtained below, confirming that the marginal worker has to be at the minimum wage.

where  $\hat{\eta}(\gamma) \equiv \max \{w_M/\phi; [\max\{w_M, \gamma\} - \delta + k^* + c]/\theta\}$ . The second term in brackets is explained by noting that workers who have high enough ability such that (8) does not hold could potentially finance their training, but in that case, they need to have sufficient funds in the first period, i.e. low enough  $\gamma$ , so that they are not affected by credit constraints. Therefore, minimum wages increase training for workers who were previously unable to finance their training, and have second period wages close to the minimum. They reduce training for workers who were financing their own training, and have second period wage sufficiently above the minimum so that it is not in the interest of their employers to provide them with training.

As the above discussion suggests  $T^{MW}$  could be smaller than or greater than  $T$ , so in this more general model, minimum wages could increase or reduce training among those still employed. More specifically, they will increase training for some workers, especially those at low wages affected by credit constraints, and reduce it for those taking wage cuts to finance their own training. This discussion also suggests that we expect  $T^{MW}$  to be greater than  $T$  (i.e. minimum wages to increase training), when most workers are credit constrained, and when there are many workers with abilities between  $\eta^*$  and  $w_M - c$ , for whom minimum wages lead to firm-sponsored training.

Finally, we have so far compared an economy without minimum wages to one with a binding minimum wage. The results of comparing two economies with different levels of minimum wages are similar. The only difference is that in a model with a binding but low minimum wage, some workers with low values of  $\gamma_i$  are already prevented from paying for their own training. So increasing an already binding, but relatively low, minimum wage is perhaps more likely to increase training than introducing one in an economy without minimum wages. Whether this happens is an empirical question, and we will now provide some new evidence on this.

## 5 Empirical Strategy and Data

The federal minimum wage was unchanged between 1981 and 1990, but various states imposed their own minima above the federal level during the late 1980s. While minimum wages were rather uniform across states before 1987 and after 1991, there was substantial dispersion between these dates. We will exploit this variation. Table 1 displays the statutory minimum wages in the US states over this period.

We use two complementary approaches to identifying the impact of minimum wages on training. The first one, which we find most compelling, looks at the training of workers who are directly affected by an increase in the state or federal minimum wage. The second approach looks at the relation between training and a measure of how binding

the minimum wage is across regions. This latter approach is most closely related to the empirical work in the previous literature, and therefore serves as a check on the robustness of our results.

## 5.1 Empirical Specification

The most direct way to estimate the impact of the minimum wage on training outcomes is to look at workers who are actually affected by changes in the minimum wage. In order to illustrate the approach, consider the following regression equation

$$\tau_{irt} = \alpha mw_{irt} + \beta' x_{it} + d_t + v_r + \mu_i + \varepsilon_{irt} \quad (12)$$

where  $\tau_{irt}$  is a measure of training for individual  $i$  in region  $r$  at time  $t$ ,  $d_t$ ,  $v_r$  and  $\mu_i$  are time, region, and individual effects, and  $x_{it}$  are other individual characteristics like education, age, gender, and information about the job an individual holds.  $mw_{irt}$  is a measure of whether the minimum wage binds for individual  $i$  in region  $r$  at time  $t$ . One measure for  $mw_{irt}$  would be whether the actual wage of an individual is close to the minimum wage for the region and time period. This has been the approach of Schiller (1994) and Grossberg and Sicilian (1999) using cross-sectional variation. But this strategy has the problem that other, possibly unobserved, person characteristics which are correlated with the individual's wage will also tend to be correlated with training receipt.

We therefore difference eq. (12) to obtain

$$\Delta\tau_{irt} = \alpha\Delta mw_{irt} + \beta'\Delta x_{it} + \Delta d_t + \Delta\varepsilon_{irt}. \quad (13)$$

Changes in training should now be related to changes in whether a worker is affected by the minimum wage. As a measure of  $\Delta mw_{irt}$  we use a dummy variable which indicates that the minimum wage increased from one year to the next, and the worker earned below the new minimum wage in the base year. This measure captures workers who are directly affected by a change in the minimum wage, similar to Card's (1992) analysis of employment effects. The measure relies purely on the variation of the minimum wage and base period wages, but not changes in individual wages, which may be correlated with when a worker received training. Our analysis will focus on individuals who do not move between states because moving would also confound  $\Delta mw_{irt}$  with behavioral effects. Therefore, differencing has eliminated the region effect in eq. (13).

Although we feel that this analysis of eq. (13) exploits variation of the minimum wage most directly, hence gives the most reliable results, previous studies have relied on eq. (12) to study the impact of the level of the minimum wage on training. To

make our results more comparable to these previous studies, we also undertake analysis of the levels equation. Rather than indicating affected workers directly, the variable  $mw_{irt}$  will in this case measure how binding the minimum wage is in a particular state or region. We start by using simply the actual minimum wage in region  $r$  at time  $t$ ,  $w_{rt}^m$ . What matters for the theory, however, is how binding the minimum wage is. The same federal minimum affects more workers in low wage regions than in high wage regions, and this source of variation would not be captured by  $w_{rt}^m$ . Therefore, we prefer specifications in which a given level of the minimum wage is allowed to have different effects in different regions depending on their average wages. For this purpose, we consider relative minimum wage measures  $w_{rt}^m/\bar{w}_r$ , where  $\bar{w}_r$  is a measure of the location of the wage distribution in region  $r$  (the median wage of older workers) over the whole sample period. This measure  $\bar{w}_r$  should not be affected by the minimum wage itself, and therefore just parameterizes the wage distribution, and  $w_{rt}^m/\bar{w}_r$  measures how high the minimum wage is relative to the region's wage distribution. Notice that we are not using  $\bar{w}_{rt}$  which would move with the business cycle at the regional level, and might create a spurious correlation if training incidence were also cyclical.

An obvious choice for the regions are states, since minimum wages vary at the state level. However, the wage distribution also varies within states, so that the relative minimum wage measure can be defined for smaller regions. Apart from states, we use two other measures. One partitions states into SMSAs and non-SMSA parts as our region definition. This lets us exploit the often substantial variation in wage levels between large SMSAs (like New York City) and rural areas in a state (like upstate New York) in the analysis. In total, we distinguish 136 regions. Details on the construction of these are available in Appendix 2. When we use states or these smaller regions, our measure for  $\bar{w}_r$  is the average of the median wages of male workers age 35-54 in each year between 1987 and 1992. The second "region" definition distinguishes between the male and female wage distribution within states, i.e.  $\bar{w}_r$  uses the average of the median wages of male workers age 35-54 in the state if the respondent is male, and of the median wages of female workers age 35-54 if the respondent is female. This measure exploits the fact that women should be more affected by a given minimum wage than men because women tend to earn less than men.

In the right hand panel of Table 1 we show the relative minimum wages using the male medians by state and year over the sample period. It is apparent that there is substantially more dispersion in this measure (which is still coarser than the other regional measures we use below) than in the minimum wage itself. For example, the federal minimum wage increases between 1989 and 1991 raised the relative minimum wage by only 0.055 in New Jersey but by 0.085 in Arkansas, both states without a state

minimum wage above the federal level in 1989. This illustrates how the scaling of the minimum wage measure leverages the increase in the federal minimum wage across states with different wage distributions.

There are a number of practical problems in implementing the estimation of equations (12) and (13). For example, training is not easily defined at a point in time. Because most training spells only last for a short period of time, we will define  $\tau_{rit}$  as referring to all incidents of training within a single year. Thus,  $\tau_{rit}$  will be 1 if the individual received any training during the year, and 0 otherwise. In the differenced equation (13), the dependent variable takes on the values -1, 0, and 1. We will estimate the models as linear probability models, facilitating differencing and the inclusion of fixed effects in the levels version.

Looking at a time period as long as a year has its drawbacks. The minimum wage may change within a year. In order to minimize the impact of this, we look at periods of 12 months starting in April and ending in March. Both federal minimum wage increases in 1990 and 1991 went into effect April 1. Some state minimum wage increases also took effect on April 1, but many did not. Whenever the minimum wage changed during the year, we use an employment weighted average of the minimum wage in effect during the year.

The covariance matrix of the error term in eq. (13) will have a first order moving average structure at the individual level. We therefore estimate standard errors with the Huber estimator, which is robust to arbitrary cluster effects at the individual level. This covariance estimator is consistent in this case but not efficient. In (12), the key regressor, the minimum wage variable, only varies at the region and year level while we use individual level data. Conventional standard errors may therefore overstate the precision of the estimates (Moulton, 1986). Suppose the error term has the form  $\varepsilon_{rit} = \lambda_i + v_{rt} + \xi_{rit}$ , i.e. the error is composed of an individual level component  $\lambda_i$ , a region\*time component  $v_{rt}$ , and a component  $\xi_{rit}$ , which is uncorrelated across individuals, regions, and time periods. Notice that the error  $\varepsilon_{rit}$  will be heteroskedastic, since we are estimating a linear probability model. There is no straightforward way to calculate consistent standard errors for this error structure. We extend the standard Huber estimator to allow for both an individual level component and a region\*time component in the error term. This estimator seems to perform well in practice in samples of our size. We report the results of some small Monte Carlo experiments in Appendix 3.

## 5.2 The Data

Our data on training come from the National Longitudinal Survey of Youth (NLSY). The NLSY is a panel of youths aged 14 to 21 in 1979. This dataset is particularly suitable for this project because it samples young workers, and it oversamples those from disadvantaged backgrounds, who are more likely to work in jobs at or slightly above the minimum wage (see Card and Krueger, 1995). We will follow the cohorts interviewed in the NLSY from 1987 to 1992, years of significant changes in state and federal minimum wages. During the 1988 to 1992 surveys, the NLSY asked a consistent set of questions about on-the-job training during the previous year. The information about the training includes length and type of the program, site, and whether the explicit costs of the training were paid for by the employer or someone else. The first set of training questions in 1988 refer to a longer time frame than the questions in subsequent years, because no similar data were collected in 1987. In 1993, the module on training in the survey was expanded substantially and the survey switched from paper and pencil to computer assisted interviewing. We use some data from the 1993 survey to complement information on training during the April 1992 to March 1993 year. There were some other minor additions to the training questions before 1993 as well.

The sequence of questions on training begins with a lead-in stating

“I would now like to ask you about other types of schooling and training you may have had, excluding regular schooling we have already talked about. Some sources of occupational training programs include government training programs, business schools, apprenticeship programs, vocational or technical institutes, correspondence courses, company or military training, seminars, and adult education courses.”

This suggests that respondents will mostly report relatively formal training programs and neglect other sources of informal on-the-job training, a suspicion which has been substantiated by Loewenstein and Spletzer (1999) using the more detailed training data in the NLSY starting in 1993. While this may be a drawback, this limitation of the data is pervasive in this literature.

We are only interested in training programs which take place in firms, or are sponsored directly by the employer, not in courses taken by individuals outside work on their own initiative or government sponsored training programs. We therefore classify the following forms of training as “employer related training”: any training for which the respondent gives as venue an apprenticeship program; formal company training; or seminars or training programs at work run by someone other than the employer; or if

the respondent classifies the training as on-the-job training or work experience; or if the employer paid for the training (even though we do not interpret the answer to this question as the employer necessarily bearing the investment costs). We do not classify a training programs as employer related if the training was partly paid for by a government program.<sup>9</sup> For each training program we record the start date as reported on the 1988 to 1993 surveys. If this date falls within an April to March period between 1987 and 1993, we assign the training to this particular year. We treat observations with a missing start date as missing.<sup>10</sup>

We do not use any information on the job or employer in the estimation directly. However, training often takes place when individuals start new jobs and minimum wage increases may affect turnover and hiring. Hence, it seems important to control for turnover in some fashion. We include a dummy variable for whether a respondent started any new job within a particular year from information in the work history module of the data.

We limit our sample to workers who have 12 years of education or less, a group most likely to be affected by the minimum wage. We use the oversamples of blacks, Hispanics and poor whites in the analysis, but we drop the military subsample. The results are weighted by the NLSY sampling weights throughout. Individuals living in the District of Columbia are excluded, because DC had a plethora of different minimum wage rules, making it hard to define a sensible overall measure of the minimum wage. Our basic sample includes all workers who report a wage at the interview for the current year and for the past year, and who were employed for at least one month during the year according to the information in the work history module.<sup>11</sup> We also restrict the sample to those with valid wage information. For the analysis using eq. (13) we also include individuals who move between states.

In addition, we use the 1987 to 1993 outgoing rotation group files of the CPS. We calculate the median wage for workers age 35 to 54 from this data source to construct the relative minimum wage measure. The CPS outgoing rotation groups are large enough to do this even for smaller areas within states. For example, each region-year cell has at least 27 observations on older male workers, but few cells are that small, and the median number of observations is 274.

---

<sup>9</sup>The most common government program, JTPA, involves wage subsidies up to 50 percent if trainees are placed with private sector employers. Hence, the incidence of training under this program is unlikely to be affected by minimum wage legislation.

<sup>10</sup>The NLSY also provides information on the length of training. We do not use the duration of training directly, because of the frequency of missing values but we checked that our results are robust to excluding very short training spells.

<sup>11</sup>The NLSY refers to this wage measure as the CPS wage because of the CPS style question. The CPS wage may refer to a prior job if the respondent is not working at the time of the interview.

## 6 Empirical Results

Table 2 reports means of some demographic indicators for the three samples from the NLSY. All samples include respondents with 12 years of schooling or less working in at least one month during the year. Since young adults still obtain additional schooling in this age range, the samples changes over time. The non-mover sample differs little from the unrestricted sample. About 20 to 25 percent of each sample report having started a new job during the calendar year. While this includes secondary and temporary jobs, the number matches relatively closely the fraction of low education workers in the same age group who report tenure of 12 month or less on their current job in the January 1991 CPS. This relatively high rate of job starts reflects the large turnover among young, low skilled workers.

The average nominal hourly wage in the basic sample rises from about \$7.50 in 1987 to almost \$10 in 1992. While a number of sample members earn the minimum wage or less, as shown in the last row, the majority of respondents earn far above the minimum wage. These samples therefore include many workers whose wage is not directly affected by the minimum. These higher wage workers effectively form our control group in the differenced analysis. When we look at the impact of the level of the minimum wage, on the other hand, most of the sample members are not directly affected by the minimum wage, a problem which has also affected previous studies. In order to address this issue we also use a low wage sample, defined as workers employed at a wage which is 150 percent of the federal minimum or less in the *previous* year. The last two columns give the basic descriptive statistics for this sample, which is about a third the size of the basic sample. These low wage workers include more women and blacks. The number of high school dropouts is only slightly higher. Average wages in this sample are much lower and do not grow substantially over the sample period, due to the sample selection. We think of this group as much more likely to be actually subject to the minimum wage or spillovers resulting from it. Since a larger fraction of these workers are directly affected by the minimum wage, according to standard theory we should find larger negative effects in this sample.

Table 3 reports sample means for some of the key variables in our regressions. The first row of the table reports the incidence of training, which is around 10 percent. The only exception is 1987, where the measured incidence is much lower. This is presumably due to the fact that the training questions were first asked on the 1988 survey so that the questions referring to 1987 had a longer recall period than for later years. There is also a small drop in 1991, possibly due to the recession. If we exclude very short training spells of 1 day or less, the incidence drops to about 7 to 8 percent.



Our measure of the minimum wage in a region is the higher of the state and the federal minimum wage. In Table 3, we report the average minimum wage in a year across all regions. These averages are weighted by the residence of our sample population. In 1987, the average minimum wage was 3.36, when only Alaska, and the New England states had state minima above the federal minimum wage of 3.35. It rises to 3.44 in 1988, mainly reflecting the increase in the California state minimum, and increases further to 3.51 in 1989, due to new minimum wage laws in other states, among them Oregon, Washington, and Pennsylvania. The standard deviation across respondents reaches a high of 29 cents in this year, indicating a substantial amount of variation in state minimum wage levels. This variation drops substantially in 1990 and 1991, when the two federal increases raise the minima in those states which had not taken action before. The averages of 3.87 and 4.26 are now only slightly above the federal minimum wages of 3.80 and 4.25. In 1992, New Jersey raises its minimum wage to 5.05, increasing the spread again.

For our analysis of equation (13), we use four different measures for workers affected by the minimum wage. The first one includes all workers who earned less than the new minimum wage in a year prior to a minimum wage increase. This includes workers who report a wage below the initial minimum. The top panel of Table 3 shows that about 1.4 percent of the sample were affected by minimum wage increases in 1988, mostly due to the increase in the California minimum wage. The state increases in 1989 affect slightly more workers, but a large fraction of workers (7 and 9 percent, respectively) is only affected by the federal increases in 1990 and 1991. The second measure excludes workers below the minimum wage in the base year. It is not completely clear whether these workers should be considered affected or not. Minimum wage coverage was fairly universal during the time period we consider, so that these reports presumably reflect mostly measurement error.<sup>12</sup> We therefore prefer the first measure which includes these workers. Excluding them cuts the fraction of affected workers about in half in each year.

An increase in the minimum may also affect the wages of higher wage workers via spillover effects, and our model then predicts that their training should be affected. Lee (1999), for example, finds large spillover effects from the minimum wage changes during our sample period, and we report similar results below in Table 5. In order to investigate whether spillovers affect our results, we also included workers who initially earned above the new minimum in the affected group. We choose alternatively 150 percent and 130 percent of the new minimum wage. This yields about two to four times as many affected workers as our original measure. The bottom panel of Table 3 reports the relative minimum wage measures, i.e. the minimum wage divided by the

---

<sup>12</sup>It is also possible that respondents receiving tips do not include these in the wage measure.

state, regional, or state/gender median for older workers. The changes in these measures over time reflect again the increases in various state minimum wages and the federal increases in 1990 and 1991.

## 6.1 Results Using Affected Workers

Regression results for the first differenced version of the model are displayed in Table 4. Apart from the minimum wage effect, the regressions include a constant (capturing any linear effects of age), a full set of time dummies and variables for the change in high school graduation status and for whether the worker took a new job. The coefficients on these covariates are quite sensible. High school graduates are 7 percentage points more likely to receive training. This effect is imprecisely estimated because there are few workers who acquire a high school degree or equivalent during the sample period. Workers starting a new job are 3 percentage points more likely to receive training and this effect is estimated quite precisely. These estimates are sensible and demonstrate that the training variable is able to pick up the expected variation in the data.

The coefficients on the variable for affected workers are directly interpretable as the effect of raising the minimum wage on the incidence of training. The four different specifications correspond to four different definitions of the “affected” variable. In column (1), we define all workers whose wage in the previous year is below the current minimum as affected. The point estimate indicates that being affected by a minimum wage increase raises the probability of receiving training by 1 percentage point. The effect is not significant, however, and a 95 percent confidence interval includes negative effects as large as 1.8 percentage points.

The specification in column (1) counts workers who were reporting wages below the previous minimum in the base year as affected. While we prefer this specification, as discussed above, it is possible that some of the wages below the initial minimum refer to uncovered workers. Therefore we exclude these workers from the affected group in the second column. The results are not very different. Neither of the previous specifications allows spillovers of the minimum wage on workers with slightly higher wages. If we redefine the affected group as workers whose previous wage was within 150 or 130 percent of the new minimum wage in columns (3) and (4), we find basically zero effects of minimum wages on training, but these specifications may count too many workers as affected.

The training question that we use includes very short training programs, and it is important to ensure that our results are not sensitive to eliminating these short spells. With this purpose, we also repeated these regressions excluding training spells lasting

a single day or less. We obtain slightly smaller coefficients for affected workers, ranging from 0.004 for the measure in column (2) to -0.013 for the measure in column (4), while the standard errors were basically unchanged.

Overall, the results contrast sharply with the predictions of the standard theory. The standard model predicts that all general training should be eliminated for workers affected by a minimum wage increase. The average incidence of training in the sample is about 10 percent, but this may not be the right baseline, since it includes higher wage workers not affected by the minimum wage. Average training incidence for workers affected by a minimum wage increase in the following year is 5.2 percent for the measure of affected workers used in column (1). The 95 percent confidence interval is consistent with declines in training as large as 1.8 percentage points. This means we can reject that the minimum wage eliminates more than a third of the training in this group. Similar conclusions are obtained for the other measures of affected workers.

Since according to the standard theory, an increase in the minimum wage should eliminate all general training among affected workers, our results could only be consistent with the standard theory if most training is specific. But Loewenstein and Spletzer (1997) find in later waves of the NLSY 63 percent of all training to be general, and in another 14 percent of training programs, most of the skills are reported to be general. They find similar results with other datasets as well. We can therefore reject the strongest implication of the standard theory that minimum wages should wipe out all training for affected workers. However, we cannot rule out that minimum wages simply reduce training for affected workers, without completely eliminating it. Nevertheless, our results indicate a substantially smaller effect of minimum wages on training than the estimates in the previous literature.<sup>13</sup>

## 6.2 Results Using Minimum Wage Changes

While the results in Table 4 indicate no adverse effects of minimum wages on training, our methodology differs somewhat from the previous literature. In this section we present alternative results based on the regression equation (12). The sample we use in this section also includes workers who move from state to state between interviews, but this sample differs little from the non-mover sample.

Recall that this sample contains many workers not directly affected by the minimum wage, so it is natural to worry about whether minimum wages will have a significant effect on the earnings of workers in this sample, since in the absence of such a finding,

---

<sup>13</sup>For example, as we discussed in Section 2, Neumark and Wascher's (1998) estimates imply that minimum wages reduce training among workers close to the minimum wage by much more than the existing amount of training.

we may expect no effect on training incidence either. Before turning to the impact of minimum wages on training, we therefore look at the effect of the minimum wage on actual wages. Table 5 displays quantile regression estimates of the real wage of workers in this sample on the real value of the minimum wage and a full set of year and state dummies. The first feature of the results is that low quantiles of the wage distribution are affected by changes in the minimum wage, which shows that minimum wages do have an effect on the wages of low paid workers. So according to the standard theory, there should be a negative effect on training.

A one dollar increase in the minimum wage raises wages of the 10<sup>th</sup> percentile worker in the NLSY by 37 cents, which may seem small. The second column in the table replicates these results with a comparable sample using the CPS outgoing rotation groups, with similar results, showing that these results are not particular to the NLSY. There are various reasons to expect why the coefficients even at the low quantiles should be less than one. There is certainly much measurement error in the wage reports of workers, biasing these coefficients down. Furthermore, Table 2 revealed that even workers at the 10<sup>th</sup> percentile will typically earn above the old minimum wage already, and therefore they may not receive the full increase when the minimum wage goes up. Nevertheless, the results in Table 5 show that workers as high as the 30<sup>th</sup> percentile of the wage distribution may be affected by minimum wage changes, and therefore their training may also be affected.

Table 5 also shows that there are many workers in this sample who are not affected by minimum wages. This means that our estimates of training effects will tend to be biased towards zero. This motivates our strategy to compare our basic results to those in the lower wage sample using only workers earning less than 150 percent of the federal minimum wage in the previous year, corresponding roughly to the workers up to the 30<sup>th</sup> percentile of the original sample. This sample more closely approximates the workers actually affected by the minimum wage or by spillovers resulting from the minimum. If the results in the larger sample are biased towards zero, then the more restrictive sample should lead to more extreme estimates.

Table 6 presents our regression results for the incidence of training on the minimum wage and relative minimum wage measures. We present four sets of regressions with and without time and region fixed effects, as well as one specification which includes individual instead of region fixed effects.<sup>14</sup> Other covariates in the regressions are dummies for blacks, Hispanics, females, high school graduates, and whether the individual

---

<sup>14</sup>If training this year makes it less likely that the individual will be trained in the future, the individual effects model may be problematic. However, in our sample, training is positively correlated over time.

started a new job during the year, and a linear variable for age.<sup>15</sup> The coefficients on the demographic covariates are again sensible. Minorities receive significantly less training. Contrary to the typical finding in the literature, women receive as much training as men in this low wage group. High school dropouts receive less training, and workers starting new jobs receive more training. Not surprisingly, this latter effect is strongest when individual effects are controlled for.

If no region or time effects are included, the effect of the minimum wage measure is positive and significant. This reflects that states with higher minimum wages tend to have more training. Because these states tend to have higher wages in general, this may simply mean that high and low wage states differ, for example in terms of their industrial and occupational structure. In fact, the positive result in column (2) vanishes when we look at the relative minimum wage measures below or when we control for state and time effects in column (4). This latter specification, which we prefer, suggests that higher minimum wages have a small negative effect on training. The federal minimum wage increased by 81 cents in real terms from 1989 to 1991. This increase led to about 0.6 percentage points less of training. We find similar results including individual fixed effects instead of state effects in column (5).<sup>16</sup>

The estimates are more negative if we look at the low wage sample in row 2. Unfortunately, these results are very imprecisely estimated and are consistent with both substantial negative and positive effects. One reason for this is that the minimum wage variable we are using in the first two rows does not exploit a relevant part of the information, which is that a given level of the minimum wage will have very different effects depending on how high wages are in a region. To exploit this information, we next turn to the relative minimum wage measures.

Rows 3 and 4 of the table report results where we divide the minimum wage by the median wage of older males in the state. The specification which exploits primarily the cross region variation in the minimum wage measure in column (2) now results in a moderate negative coefficient. In row 5 below, where we use more detailed regions and obtain more precise results, this coefficient is even significant. This specification compares most closely to the approach of Leighton and Mincer (1981). But the results in columns (4) or (5), which are much closer to zero, indicate again that the minimum wage measure was picking up across region differences in training incidence. We feel

---

<sup>15</sup>Curvature of the age profile is empirically unimportant for the small age range in the sample, therefore we do not include higher order terms.

<sup>16</sup>In the specification with individual fixed effects, only the region\*time clusters remain in the error term. The standard Huber estimator allowing for these clusters is consistent. Our standard errors in column (5) are therefore the most reliable. Interestingly, the standard errors on the minimum wage variable do not differ substantially from those in column (4).

that the estimates including time and region or individual effects are most reliable. Some of the time variation of our training measure is likely due to the change in the survey questions and cyclical variations. Differences in training incidence across regions may reflect differences in industrial and occupational compositions. The negative effect of the minimum wage in column (2) is therefore likely to reflect the fact that higher wage regions tend to have more training. The results including time and region effects, which are the most reliable, do not indicate any negative effects of minimum wages on training.<sup>17</sup>

In order to interpret the magnitudes of the estimates and confidence intervals using this relative measure of the minimum wage, return to the bottom panel of Table 3. The relative minimum wages increased by about 5 to 6 percentage points from 1989 to 1991. To gauge the impact of the federal increases, it is more useful to calculate the means only for the states that were subject to the federal minimum in 1989. For these states, the federal increases raised the relative minimum wage measures by 0.07 using the male state or region medians, and by 0.08 using the state/gender medians. This means that a 95 percent confidence interval for the estimates in column (4) using state medians for the unrestricted sample excludes negative changes in training of 2.9 percentage points or larger in response to the federal increases. It is instructive to compare this result to the findings by Neumark and Wascher (1998), for example. Their point estimates imply that the California state minimum wage (which was similar in magnitude to the federal increase between 1989 and 1991) led to a decline in formal training of 3.2 percentage points among young workers age 20-24, a group with a similar average training incidence as our sample. We can reject a decrease of this size for our sample.

Looking at the low wage sample, we obtain a virtually identical point estimate when we include time and state effects. When we include individual instead of state effects, the point estimate is actually positive now. Thus, the results with the low wage sample do not indicate that the estimates for the basic sample were attenuated. These conclusions are unchanged when we look at the minimum wage measure scaled by alternative values of the median wage. In rows 5 and 6, we use the median wage for 136 regions smaller than the states. We find small positive point estimates for the basic sample, and slightly larger estimates in the low wage sample. To the degree that these results indicate attenuation in the basic sample, they suggest that the actual effect of minimum wages on training is positive, rather than negative. The within estimates using these smaller regions turn out not to be any more precise than the state results, however.

---

<sup>17</sup>Our theory predicts that minimum wages reduce employment, but our samples focus on employed workers. This might bias our estimates up. When we use a sample of respondents including non-workers, we actually find slightly more positive results, however.

The last two rows in Table 7 present results scaling the minimum wage by the median wage for older workers in the same state and of the same gender as the respondent. The results for the basic sample are again rather similar to the previous measures. The results are now somewhat more precise, because there is more variation in the measure exploiting gender differences. For the low wage sample, the results are a bit different. Controlling for time and state effects, we now find a negative effect for minimum wages. However, this result is not replicated when we control for individual effects instead. In addition, if we tighten the low wage sample further to include only workers who earned less than 130 percent of the federal minimum wage in the previous year, we find a coefficient of 0.097 (with a standard error of 0.294). We conclude that the negative estimate for the low wage sample is much more likely an indication of the sampling variation of these estimates rather than evidence that the estimates for the basic sample are attenuated.

## 7 Concluding Remarks

Standard theory of human capital makes an unambiguous prediction regarding the impact of minimum wages: training should decline for affected workers, because workers finance the provision of the general component of training by taking lower wages, and minimum wages prevent this. We show that this conclusion is not robust. In labor markets where employers obtain rents from the employment relationship, the theory has quite different implications regarding the impact of minimum wages on training. While minimum wages reduce training for some workers as in the standard theory, they also encourage firm-sponsored training and increase human capital investments for other workers. Therefore, our model predicts offsetting effects of minimum wages on training, and is even consistent with minimum wages increasing training investments.

In the second part of the paper, we provided some new empirical evidence on the impact of minimum wages on training. Focusing on workers actually affected by minimum wage changes we find no evidence that minimum wages reduce training. In addition, we can reject negative estimates of the magnitude previously found in the literature. This result is not an artefact of our methodology, since we find comparable results using specifications which are more similar to previous studies. We obtain the most negative estimates when we simply regress training on the real minimum wage within states. However, these estimates are also the most imprecise, and therefore the least informative. When we rely on relative minimum wage measures, thus better exploiting the differential impact of the changes in the federal minimum wage, we find no negative effects on training. We probe these results by comparing estimates from a larger base-

line sample to those from a smaller sample of low wage workers. The baseline sample resembles what other researchers have used, in that it includes many workers unaffected by the minimum wage. If these unaffected workers simply obscure the impact of the minimum wage on training, we should find systematically more negative effects using low wage workers, but this is not the case, once again suggesting that binding minimum wages do not reduce training.

A drawback of our analysis is that we only focus on rather formal training. Possibly a much larger part of on-the-job learning in low wage jobs is more informal. It is plausible that formal and informal training are positively related, but our analysis does not shed any direct light on the receipt of less formal training. The NLSY introduced more extensive questions on informal training starting in 1993. Unfortunately, this is a period with little dispersion in minimum wage changes, and the NLSY sample is also aging so that fewer workers receive the minimum in the 1990s.



## 8 Appendix 1: First Period Wages With Worker Heterogeneity

First, take a  $\eta_i$  satisfying  $w_M > \phi\eta_i$ , then  $\Pi_j(\eta_i) = \Pi_j = k^* - k_j$  solves for

$$w_1(\eta_i, T) = \max \{w_M, (\phi + \theta)\eta_i + \delta - w_M - k^* - c\}. \quad (14)$$

So (14) gives first period wages when the minimum wage is binding.

In contrast, we could have  $w_M \in (\eta_i - c, \phi\eta_i)$ , that is the minimum wage is binding in the second period if worker  $i$  is not trained, but does not bind when he is trained. Then,  $\Pi_j(\eta_i) = \Pi_j = k^* - k_j$  solves for

$$w_1(\eta_i, T) = \max \{w_M, \theta\eta_i + \delta - k^* - c\}.$$

Next, if  $\eta_i$  satisfies (9), then by construction, worker  $i$  takes a wage cut in the first period to obtain training, so:

$$w_1(\eta_i, T) = \theta\eta_i + \delta - k^* - c > w_M.$$

Finally, if (8) and (9) do not hold, then,

$$w_1(\eta_i, NT) = \max \{w_M, (1 + \theta)\eta_i + \delta - k^*\},$$

which completes the characterization of first period wages for given  $k^*$ .

## 9 Appendix 2: Construction of Regions

Some of the minimum wage measures used in this paper scale the state minimum wage level by a regional reference wage. An NLSY respondent's reference wage is the median wage in his/her geographic region for 35-54 year old male workers calculated from the CPS Outgoing Rotation Groups. This appendix explains how the regions were constructed.

To account for the fact that there is significant variation in average wages within many states, and to allow the minimum wage measure to have greater within-state variation, we used information on an NLSY respondent's CMSA/MSA of residence, ("msa\_code"). Because of sample sizes, we used msa\_code only for individuals residing in the large MSAs. For all other areas (i.e., smaller MSAs and non-MSAs), we used data on whether an individual resided in an MSA ("msa\_status").

The following list contains the 35 MSAs we classified as large. We included the 20 CMSAs defined by the 1990 FIPS Guidelines (plus the St. Louis, MO CMSA, while we

do not distinguish the Providence, RI CMSA from other metropolitan parts of RI) and added 15 MSAs with large populations.

Atlanta, GA  
Baltimore, MD  
Boston-Lawrence-Salem, MA-NH CMSA  
Buffalo-Niagara Falls, NY CMSA  
Chicago-Gary Lake County, IL-IN CMSA  
Cincinnati-Hamilton, OH-KY CMSA  
Cleveland-Akron-Lorraine, OH CMSA  
Columbus, OH  
Dallas-Fort Worth, TX CMSA  
Denver-Boulder, CO CMSA  
Detroit-Ann Arbor, MI CMSA  
Hartford-New Britain-Middletown, CT CMSA  
Houston, TX CMSA  
Indianapolis, IN  
Kansas City, KS  
Los Angeles-Anaheim-Riverside, CA CMSA  
Miami-Ft. Lauderdale, FL CMSA  
Milwaukee-Racine, WI CMSA  
Minneapolis-St. Paul, MN  
New Orleans, LA  
New York-New Jersey-Long Island, NY-NJ-CT CMSA  
Norfolk-Virginia Beach-Newport News, VA  
Orlando, FL  
Philadelphia-Wilmington-Trenton, PA-DE-NJ CMSA  
Phoenix , AZ  
Pittsburgh-Beaver Valley, PA CMSA  
Portland-Vancouver, OR CMSA  
Sacramento, CA  
St. Louis, MO CMSA  
San Antonio, TX  
San Diego, CA  
San Francisco-Oakland-San Jose, CA CMSA  
Seattle-Tacoma, WA CMSA  
Tampa-St. Petersburg-Clearwater, FL  
Washington, DC, DC-MD-VA

The full set of regions consists of state-`msa_code` level regions for the above 35 MSAs and state-`msa_status` level regions for all other areas. The example of an individual living in Massachusetts illustrates how the reference wage is chosen. If the individual lived in the Boston CMSA, the reference wage would be the median wage for 35-54 year old male workers living in the Massachusetts part of the Boston CMSA (the “MSA median”). If the individual lived in a metropolitan area other than the Boston CMSA, the reference wage would be the median wage for 35-54 year old male workers living in Massachusetts metropolitan areas other than the Boston CMSA. If the individual lived in a non-metropolitan area, the reference wage would be the median wage for 35-54 year old male workers living in Massachusetts non-metropolitan areas (“non-MSA median”).

For confidentiality reasons, both the NLSY and CPS sometimes suppress the specific MSA of residence for individuals in smaller MSAs and/or smaller states. In the CPS data, there were five states in which `msa_status` was sometimes suppressed: Maryland, Nevada, Rhode Island, Utah and Wyoming. We classified these suppressed observations as non-MSA, with the following rationale. In Maryland, the suppressed observations are from the Maryland part of the Philadelphia CMSA. In Nevada, they are from the non-central city part of the Reno MSA. In Rhode Island, they are from the Rhode Island part of the New London MSA. In Utah, they are from the Utah part of the Flagstaff MSA. Finally, no MSAs in Wyoming are identified. For all five states, an individual was either classified as living in an MSA or the information was suppressed, i.e. observations from fringes of MSAs were grouped with the non-MSA observations to ensure confidentiality. Thus, classifying all the individuals living in areas with suppressed codes as non-MSA seems the most sensible choice.

In addition, there are MSAs which straddle different states and sometimes part of an MSA in a particular state is not identified if the part is too small. This was the case for: the Indiana part of the Cincinnati MSA, the Maryland part of the Philadelphia CMSA, the North Carolina part of the Norfolk MSA and the Wisconsin part of the Chicago CMSA and Minneapolis MSA. The CPS groups these areas with other metropolitan areas, and thus it is appropriate to use the state-MSA median for these observations in the NLSY. Finally, in Maryland, the Cumberland, MD-WV MSA and the Hagerstown, MD MSA are not separately identified. Here we assign the non-MSA median since Maryland’s non-MSA median is really a median excluding the Baltimore and Washington, D.C. MSAs.

In the NLSY data, there were two complications. First, as in the CPS, `msa_status` was suppressed for certain individuals. We dropped the suppressed observations (62 out of 21,680 observations from the Basic Sample). Second, the NLSY uses NECMAs for the New England states whereas the CPS uses CMSA/MSAs. The NLSY switched to

NECMA codes in 1988; we constructed the NECMA codes for 1987 by using information on state and county of residence. We then mapped NECMA codes into MSA codes. Specifically, individuals living in the Boston NECMA or Hartford NECMA were classified as living in the Boston CMSA and Hartford CMSA, respectively. Also, individuals living in any NECMA were classified as living in an MSA.

All together, the unrestricted NLSY sample contains 21,618 observations from 136 regions and 753 region-year categories. Of the 136 regions, 44 correspond to state-msa.code level regions, 45 to state-MSA level regions and 47 to state-non-MSA regions. These regions form a partition of the country. The median number of observations in each region-year cell in the CPS is 274 and the range is 27 (for the Indiana part of the Chicago CMSA in 1991) to 2261 (the Los Angeles CMSA in 1990). Less than 10 percent of all cells have fewer than a 100 observations and less than 5 percent of the NLSY observations fall into those cells.

## 10 Appendix 3: Estimation of the Standard Errors

We assume that the error term has the form  $\varepsilon_{ij} = \lambda_i + v_j + \xi_{ij}$  where  $i$  denotes individual and  $j$  denotes region\*time. Let  $x_{ij}$  be a vector of right-hand side variables,

$$X_i = \begin{bmatrix} x'_{i1} \\ x'_{i2} \\ \dots \\ x'_{iJ} \end{bmatrix}$$

the matrix of right-hand side variables for individual  $i$ , and  $X$  the stacked matrix of all the  $X_i$ 's. To simultaneously adjust for individual and region\*time effects, we use the covariance matrix

$$\hat{V} = (X'X)^{-1}X'\Omega X(X'X)^{-1}$$

where

$$X'\Omega X = \sum_i X'_i \hat{\varepsilon}_i \hat{\varepsilon}'_i X_i + \sum_{i \neq k} \sum_j \hat{\varepsilon}_{ij} \hat{\varepsilon}_{kj} x_{ij} x'_{kj}.$$

In order to get an idea how this covariance estimator performs in our sample, we conducted a small Monte Carlo experiment. For this experiment we generated samples with the same number of observations, individuals, regions, and time periods as in our unrestricted sample according to the design

$$\begin{aligned} y_{ij}^* &= 0.4 + 0.1x_{1ij} + 0.1x_{2j} + 0.1x_{3i} + \lambda_i + v_j + \xi_{ij} \\ y_{ij} &= \begin{cases} 1 & \text{if } p \leq y_{ij}^* \\ 0 & \text{if } p > y_{ij}^* \end{cases} \end{aligned}$$

where  $p$  and each of the  $x$ 's are drawn from a uniform (0,1) distribution and  $\lambda_i, v_j, \xi_{ij} \sim N(0,0.1)$ . In regression 1, we computed the OLS estimate of  $y$  on the three  $x$ 's and constructed the covariance matrix above. In regression 2, we estimated the regression of  $y$  on the  $x$ 's also including region and time fixed effects. We replicated each regression 10000 times with the following results:

regressor	Regression 1			Regression 2		
	$x_{1ij}$	$x_{2j}$	$x_{3i}$	$x_{1ij}$	$x_{2j}$	$x_{3i}$
standard deviation of $\hat{\beta}$	0.0116	0.0320	0.0126	0.0116	0.0355	0.0124
mean of estimated standard error	0.0116	0.0314	0.0124	0.0115	0.0304	0.0122
Year Effects		No			Yes	
Region Effects		No			Yes	

Regression 1 indicates that the covariance estimator does a good job in estimating the sampling variation for all regressors. However, in regression 2, the sampling variation for  $x_{2j}$ , which only varies at the region\*time level, i.e. the analogue to our minimum wage measure, is understated by about 15 percent. Of course, these results are only suggestive. In our design,  $v_j$  contributes a third to the total variance of the error term. The results would change as we change the relative variances of this error component.

## References

- [1] Acemoglu, Daron and Jörn-Steffen Pischke (1999a) “Beyond Becker: Training in imperfect labor markets,” *Economic Journal Features* 109, F112-142.
- [2] Acemoglu, Daron and Jörn-Steffen Pischke (1999b) “The structure of wages and investment in general training,” *Journal of Political Economy* 107, 539-572.
- [3] Becker, Gary (1964) *Human capital*. Chicago: The University of Chicago Press.
- [4] Ben-Porath, Yoram (1967) “The production of human capital over the life cycle,” *Journal of Political Economy* 75, 352-365.
- [5] Card, David (1992) “Using regional variation in wages to measure the effect of the federal minimum wage,” *Industrial and Labor Relations Review* 46, 22-37.
- [6] Card, David and Alan B. Krueger (1995) *Myth and measurement. The new economics of the minimum wage*. Princeton: Princeton University Press.
- [7] DiNardo, John, Nicole Fortin and Thomas Lemieux (1996) “Labor market institutions, and the distribution of wages, 1973-1992: A semiparametric approach,” *Econometrica* 64, 1001-1044.
- [8] Grossberg, Adam J. and Paul Sicilian (1999) “Minimum wages, on-the-job-training, and wage growth,” *Southern Economic Journal* 65, 539-556.
- [9] Hashimoto, Masanori (1982) “Minimum wage effects on training on the job,” *American Economic Review* 72, 1070-1087.
- [10] Lee, David (1999) “Wage inequality in the U.S. during the 1980s: Rising dispersion or falling minimum wage?” *Quarterly Journal of Economics* 114, forthcoming.
- [11] Leighton, Linda and Jacob Mincer (1981) “The effects of the minimum wage on human capital formation,” in Simon Rottenberg (ed.) *The economics of legal minimum wages*. Washington: American Enterprise Institute for Public Policy Research, 155-173.
- [12] Loewenstein, Mark A. and James R. Spletzer (1997) “General and specific training: Evidence from the NLSY,” mimeographed, Bureau of Labor Statistics.
- [13] Loewenstein, Mark A. and James R. Spletzer (1999) “Formal and informal training: Evidence from the NLSY,” in Solomon Polachek (ed.) *Research in labor economics*, vol. 18, Greenwich, CT: JAI Press.

- [14] Mincer, Jacob (1974) *Schooling, Experience, and Earnings*. New York: Columbia University Press.
- [15] Moulton, Brent R. (1986) "Random group effects and the precision of regression estimates," *Journal of Econometrics* 32, 385-397.
- [16] Neumark, David and William Wascher (1998) "Minimum wages and training revisited," NBER Working Paper No. 6651.
- [17] Rosen, Sherwin (1972) "Learning and experience in the labor market," *Journal of Human Resources* 7, 326-342.
- [18] Schiller, Bradley R. (1994) "Moving up: The training and wage gains of minimum-wage entrants," *Social Science Quarterly* 75, 622-636

</ref\_section>



**Table 1**  
**Minimum Wage and Relative Minimum by State and Year**

State	Statutory Minimum Wage						Minimum Wage/Avg. Median Wage Age 35-54					
	1987	1988	1989	1990	1991	1992	1987	1988	1989	1990	1991	1992
Alabama	3.35	3.35	3.35	3.80	4.25	4.25	0.308	0.306	0.305	0.344	0.383	0.381
Alaska	3.85	3.85	3.85	4.30	4.75	4.75	0.215	0.214	0.213	0.237	0.260	0.259
Arizona	3.35	3.35	3.35	3.80	4.25	4.25	0.262	0.261	0.260	0.293	0.326	0.325
Arkansas	3.35	3.35	3.35	3.80	4.25	4.25	0.342	0.341	0.339	0.382	0.426	0.424
California	3.35	3.35/4.25	4.25	4.25	4.25	4.25	0.234	0.279	0.294	0.292	0.291	0.290
Colorado	3.35	3.35	3.35	3.80	4.25	4.25	0.245	0.244	0.243	0.274	0.305	0.304
Connecticut	3.37/3.75	3.75/4.25	4.25	4.25	4.27	4.27	0.236	0.265	0.279	0.277	0.277	0.276
Delaware	3.35	3.35	3.35	3.80	4.25	4.25	0.242	0.241	0.240	0.270	0.301	0.300
Florida	3.35	3.35	3.35	3.80	4.25	4.25	0.304	0.303	0.302	0.340	0.378	0.377
Georgia	3.35	3.35	3.35	3.80	4.25	4.25	0.290	0.288	0.287	0.324	0.360	0.359
Hawaii	3.35/3.85	3.85	3.85	3.85	4.25	4.75/5.25	0.251	0.276	0.275	0.273	0.300	0.343
Idaho	3.35	3.35	3.35	3.80	4.25	4.25	0.296	0.294	0.293	0.330	0.368	0.366
Illinois	3.35	3.35	3.35	3.80	4.25	4.25	0.239	0.238	0.237	0.268	0.298	0.297
Indiana	3.35	3.35	3.35	3.80	4.25	4.25	0.281	0.280	0.279	0.314	0.350	0.348
Iowa	3.35	3.35	3.35/3.85	3.85/4.25	4.25/4.65	4.65	0.285	0.283	0.293	0.330	0.363	0.386
Kansas	3.35	3.35	3.35	3.80	4.25	4.25	0.270	0.269	0.267	0.301	0.335	0.334
Kentucky	3.35	3.35	3.35	3.80	4.25	4.25	0.291	0.290	0.288	0.325	0.362	0.361
Louisiana	3.35	3.35	3.35	3.80	4.25	4.25	0.289	0.288	0.286	0.323	0.359	0.358
Maryland	3.35	3.35	3.35	3.80	4.25	4.25	0.235	0.234	0.233	0.262	0.292	0.291
Massachusetts	3.55/3.65	3.65/3.75	3.75	3.80	4.25	4.25	0.250	0.256	0.256	0.258	0.287	0.286
Michigan	3.35	3.35	3.35	3.80	4.25	4.25	0.231	0.230	0.229	0.258	0.287	0.286
Minnesota	3.35/3.55	3.55/3.85	3.85/3.95	3.95/4.25 <sup>1</sup>	4.25	4.25	0.256	0.272	0.288	0.292	0.312	0.311
Mississippi	3.35	3.35	3.35	3.80	4.25	4.25	0.340	0.339	0.337	0.380	0.423	0.422
Missouri	3.35	3.35	3.35	3.80	4.25	4.25	0.279	0.278	0.277	0.312	0.348	0.346
Montana	3.35	3.35	3.35	3.80	4.25	4.25	0.292	0.290	0.289	0.326	0.363	0.361
Nebraska	3.35	3.35	3.35	3.80	4.25	4.25	0.303	0.302	0.300	0.338	0.377	0.375
Nevada	3.35	3.35	3.35	3.80	4.25	4.25	0.266	0.265	0.264	0.297	0.331	0.330
New Hampshire	3.45/3.55	3.55/3.65	3.65/3.75	3.75/3.85	4.25	4.25	0.257	0.263	0.269	0.277	0.308	0.307
New Jersey	3.35	3.35	3.35	3.80	4.25	5.05	0.216	0.216	0.214	0.242	0.269	0.319
New Mexico	3.35	3.35	3.35	3.80	4.25	4.25	0.284	0.283	0.281	0.317	0.353	0.352
New York	3.35	3.35	3.35	3.80	4.25	4.25	0.238	0.237	0.236	0.266	0.296	0.295
North Carolina	3.35	3.35	3.35	3.80	4.25	4.25	0.315	0.313	0.312	0.352	0.391	0.390
North Dakota	3.35	3.35	3.35	3.80	4.25	4.25	0.291	0.290	0.288	0.325	0.361	0.360
Ohio	3.35	3.35	3.35	3.80	4.25	4.25	0.260	0.259	0.258	0.291	0.324	0.322
Oklahoma	3.35	3.35	3.35	3.80	4.25	4.25	0.282	0.281	0.279	0.315	0.350	0.349
Oregon	3.35	3.35	3.35/4.25 <sup>2</sup>	4.25/4.75	4.75	4.75	0.262	0.261	0.287	0.336	0.364	0.363
Pennsylvania	3.35	3.35/3.70	3.70	3.80	4.25	4.25	0.269	0.273	0.294	0.300	0.334	0.333
Rhode Island	3.55/3.65	3.65/4.00	4.00/4.25	4.25	4.45	4.45	0.281	0.302	0.321	0.324	0.338	0.337
South Carolina	3.35	3.35	3.35	3.80	4.25	4.25	0.305	0.304	0.302	0.341	0.379	0.378
South Dakota	3.35	3.35	3.35	3.80	4.25	4.25	0.337	0.336	0.334	0.377	0.420	0.418
Tennessee	3.35	3.35	3.35	3.80	4.25	4.25	0.331	0.329	0.328	0.370	0.411	0.410
Texas	3.35	3.35	3.35	3.80	4.25	4.25	0.280	0.279	0.278	0.313	0.348	0.347
Utah	3.35	3.35	3.35	3.80	4.25	4.25	0.259	0.258	0.256	0.289	0.322	0.320
Vermont	3.45/3.55	3.55/3.65	3.65/3.75	3.85	4.25	4.25	0.306	0.313	0.320	0.324	0.361	0.360
Virginia	3.35	3.35	3.35	3.80	4.25	4.25	0.252	0.251	0.249	0.281	0.313	0.312
Washington	3.35	3.35/3.85	3.85/4.25	4.25	4.25	4.25	0.236	0.244	0.277	0.294	0.293	0.292
West Virginia	3.35	3.35	3.35	3.80	4.25	4.25	0.294	0.293	0.292	0.329	0.366	0.365
Wisconsin	3.35	3.35	3.35/3.65	3.80	4.25	4.25	0.270	0.269	0.286	0.302	0.336	0.335
Wyoming	3.35	3.35	3.35	3.80	4.25	4.25	0.263	0.262	0.261	0.294	0.327	0.326

<sup>1</sup> Minnesota: Only large employers covered by the new 1991 minimum wage.

<sup>2</sup> Oregon: Minimum wage changed from 3.35 to 3.85 and then 4.25 during the year.

Notes: Left hand panel of the table shows the higher of the state or federal minimum wages in each state and year. Years begin in April of the year shown until March of the following year. Multiple minima are shown if the state minimum changed during the April to March period. The right hand panel of the table shows the minimum wage divided by the average median wage for male workers 35-54 years old in the state during the 1987 to 1992 period.

**Table 2**  
**Sample Means of Demographics**  
**(Standard Deviations in Parentheses)**

<i>Variable</i>	<i>Non-mover Sample</i>		<i>Unrestricted Sample</i>		<i>Low Wage Sample</i>	
	<i>1988</i>	<i>1992</i>	<i>1987</i>	<i>1992</i>	<i>1987</i>	<i>1992</i>
Female	0.434	0.428	0.432	0.427	0.527	0.602
Black	0.134	0.132	0.128	0.133	0.169	0.203
Hispanic	0.067	0.067	0.064	0.067	0.061	0.073
Age	27.2	31.2	26.2	31.2	25.8	31.1
Less than High School	0.187	0.171	0.192	0.173	0.262	0.240
New Job	0.273	0.216	0.268	0.220	0.338	0.323
Nominal Hourly Wage	8.16	9.71	7.58	9.70	5.12	5.91
	(5.36)	(6.54)	(5.51)	(6.52)	(2.57)	(2.82)
Real Hourly Wage (1982-84 \$)	7.30	8.52	6.81	8.51	4.60	5.18
	(4.79)	(5.73)	(4.95)	(5.72)	(2.30)	(2.47)
Fraction Earning Minimum or Less	0.042	0.058	0.068	0.059	0.155	0.162
Number of Observations	3872	3094	3979	3136	1673	1048

Notes: Unbalanced panel from the NLSY. Unrestricted sample consists of individual-year observations that have high school education or less, work in at least one month of the year and in both the prior and current year have non-missing wage data. Non-mover sample excludes from the unrestricted sample individuals who have moved to a new state since the previous year. Low wage sample imposes on the unrestricted sample the restriction that the “CPS wage” in the previous year is less than or equal to 150% of the federal minimum wage. Year refers to April to March (of following calendar year). New job refers to the start of any new job during the year. Hourly wage is the “CPS wage” for workers employed at a “CPS job” only. NLSY data include black, Hispanic and poor white oversamples. The poor white oversamples were discontinued in the 1991 survey year, accounting for the lower number of observations in 1992. Statistics are weighted by the NLSY sampling weights.

**Table 3**  
**Sample Means of Key Variables**  
**(Standard Deviations in Parentheses)**

<i>Variable</i>	<i>1987</i>	<i>1988</i>	<i>1989</i>	<i>1990</i>	<i>1991</i>	<i>1992</i>
<i>Non-mover Sample</i>						
Training Incidence		0.099	0.099	0.106	0.094	0.103
Nominal Minimum Wage		3.44 (0.21)	3.51 (0.29)	3.87 (0.16)	4.26 (0.05)	4.28 (0.15)
Minimum wage increased and wage in prior year is below the current minimum wage		0.014	0.020	0.069	0.091	0.001
Minimum increased and wage in prior year is below the current minimum and above prior year minimum		0.007	0.008	0.034	0.050	0.034
Minimum wage increased and wage in prior year is below 150 % of the current minimum wage		0.062	0.085	0.263	0.293	0.011
Minimum wage increased and wage in prior year is below 130% of the current minimum wage		0.041	0.058	0.166	0.204	0.007
Number of Observations		3872	3793	3187	3122	3094
<i>Unrestricted Sample</i>						
Training Incidence	0.064	0.099	0.100	0.107	0.094	0.102
Nominal Minimum Wage	3.36 (0.06)	3.44 (0.21)	3.51 (0.29)	3.87 (0.16)	4.26 (0.05)	4.29 (0.15)
Real Minimum Wage (1982-84 \$)	3.02 (0.05)	3.08 (0.19)	3.12 (0.26)	3.42 (0.14)	3.75 (0.05)	3.76 (0.13)
Minimum Wage/Median Wage (men 35-54, states)	0.267 (0.030)	0.271 (0.027)	0.276 (0.028)	0.301 (0.031)	0.331 (0.038)	0.332 (0.037)
Minimum Wage/Median Wage (men 35-54, regions)	0.272 (0.043)	0.277 (0.041)	0.282 (0.042)	0.307 (0.045)	0.338 (0.053)	0.340 (0.051)
Minimum Wage/Median Wage (35- 54, state and gender)	0.324 (0.077)	0.330 (0.076)	0.336 (0.078)	0.366 (0.086)	0.402 (0.097)	0.404 (0.097)
Fraction Earning Minimum or Less	0.068	0.044	0.050	0.047	0.072	0.059
Number of Observations	3979	4043	3958	3295	3207	3136

Notes: See Table 2 for a description of the sample composition. Training incidence refers to employment related training. Minimum wage refers to the higher of the state or federal minimum applicable to an NLSY respondent. Median wage for 35-54 year old workers are calculated from the CPS Outgoing Rotation Groups. The poor white oversamples were discontinued in the 1991 survey year, accounting for the lower number of observations in 1990 to 1992. Statistics are weighted by the NLSY sampling weights.

**Table 4**  
**The Effect of Minimum Wage Increases on Affected Workers**

<i>Independent Variable</i>	<i>(1)</i>	<i>(2)</i>	<i>(3)</i>	<i>(4)</i>
Minimum wage increased and wage in prior year is below the current minimum wage	0.010 (0.014)	--	--	--
Minimum increased and wage in prior year is below the current minimum and above prior year minimum	--	0.016 (0.019)	--	--
Minimum wage increased and wage in prior year is below 150 % of the current minimum wage	--	--	0.005 (0.008)	--
Minimum wage increased and wage in prior year is below 130% of the current minimum wage	--	--	--	-0.002 (0.010)
Change in high school graduation status	0.070 (0.053)	0.070 (0.054)	0.070 (0.054)	0.070 (0.054)
Change in new job status	0.032 (0.008)	0.032 (0.008)	0.032 (0.008)	0.032 (0.008)

Notes: Non-mover sample, consisting of all workers with a high school education or less, who do not move between states from one year to the next. Dependent variable is the change in training incidence between two consecutive years. All regressions also include a constant and year dummies. Regressions are weighted by NSLY sampling weights. Standard errors are adjusted for the presence of individual effects in the error term, and therefore robust to the MA structure of the error. Sample size is 17074.

**Table 5**  
**Quantile Regressions for Real Wage on Real Minimum Wage**  
**Unrestricted Sample**

<i>Quantile</i>	<i>NLSY</i>	<i>CPS</i>
0.10	0.376 (0.087)	0.247 (0.024) [0.084]
0.20	0.156 (0.105)	0.137 (0.011) [0.109]
0.30	0.146 (0.163)	0.052 (0.016) [0.121]
0.40	0.049 (0.158)	0.032 (0.045) [0.102]
0.50	0.137 (0.178)	0.000 (0.025) [0.111]
0.60	0.029 (0.173)	0.130 (0.053) [0.113]
0.70	-0.120 (0.247)	0.030 (0.043) [0.135]
0.80	0.296 (0.446)	0.011 (0.044) [0.198]
0.90	0.312 (0.593)	0.157 (0.143) [0.282]
Year Effects	Yes	Yes
State Effects	Yes	Yes

Notes: Samples include respondents with a high school degree or less. Dependent variable is the real hourly wage. NLSY regressions are weighted by the NLSY sampling weights. Standard errors in parentheses are not adjusted for clusters in the error term. Bootstrapped standard errors using state\*year blocks are in brackets (100 replications). Number of observations is 21618 in the NLSY, and 119464 in the CPS.

**Table 6**  
**Panel Regressions for Training Incidence**

<i>Sample</i>	<i>Independent Variable</i>	(1)	(2)	(3)	(4)	(5)
Unrestricted	Real Minimum Wage	0.038 (0.012)	0.053 (0.019)	0.026 (0.010)	-0.007 (0.017)	-0.009 (0.020)
Low Wage	Real Minimum Wage	-0.002 (0.015)	-0.003 (0.030)	-0.001 (0.013)	-0.021 (0.032)	-0.033 (0.054)
Unrestricted	Real Minimum Wage/Avg. Median Wage (men 35-54, states)	-0.007 (0.086)	-0.189 (0.108)	0.285 (0.111)	-0.048 (0.187)	-0.080 (0.197)
Low Wage	Real Minimum Wage/Avg. Median Wage (men 35-54, states)	-0.021 (0.119)	-0.021 (0.150)	0.001 (0.147)	-0.057 (0.396)	0.201 (0.427)
Unrestricted	Real Minimum Wage/Avg. Median Wage (men 35-54, regions)	-0.075 (0.059)	-0.182 (0.067)	0.302 (0.100)	0.036 (0.192)	0.051 (0.172)
Low Wage	Real Minimum Wage/Avg. Median Wage (men 35-54, regions)	-0.053 (0.080)	-0.065 (0.079)	0.055 (0.146)	0.131 (0.372)	0.272 (0.329)
Unrestricted	Real Minimum Wage/Avg. Median Wage (35-54, state and gender)	-0.026 (0.063)	-0.171 (0.081)	0.237 (0.083)	0.048 (0.147)	-0.005 (0.166)
Low Wage	Real Minimum Wage/Avg. Median Wage (35-54, state and gender)	-0.038 (0.082)	-0.051 (0.096)	-0.009 (0.109)	-0.128 (0.248)	0.129 (0.313)
	Year Effects	No	Yes	No	Yes	Yes
	State or Region Effects	No	No	Yes	Yes	No
	Individual Effects	No	No	No	No	Yes

Notes: Sample includes respondents with a high school degree or less. Dependent variable is training incidence. All regressions also include a linear term in age, dummy variables for blacks, Hispanics, females, less than high school, whether the respondent started a new job within the year, and a constant where applicable. Regressions are weighted by the NLSY sampling weights. Standard errors in columns (1) to (4) are adjusted for the presence of state\*time or region\*time and individual effects in the error term (see Appendix 2 for details). Standard errors in columns (5) are adjusted for the presence of region\*time effects in the error term. Number of observations is 21618 in the unrestricted sample and 7026 in the low wage sample.