

IZA DP No. 6662

**High Wage Workers Match with High Wage Firms:  
Clear Evidence of the Effects of Limited Mobility Bias**

Martyn J. Andrews  
Leonard Gill  
Thorsten Schank  
Richard Upward

June 2012

# High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias

**Martyn J. Andrews**  
*University of Manchester*

**Leonard Gill**  
*University of Manchester*

**Thorsten Schank**  
*Johannes Gutenberg University of Mainz  
and IZA*

**Richard Upward**  
*University of Nottingham*

Discussion Paper No. 6662  
June 2012

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## **ABSTRACT**

### **High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias**

Positive assortative matching implies that high productivity workers and firms match together. However, there is almost no evidence of a positive correlation between the worker and firm contributions in two-way fixed-effects wage equations. This could be the result of a bias caused by standard estimation error. Using German social security records we show that the effect of this bias is substantial in samples with limited inter-firm movement. The correlation between worker and firm contributions to wage equations is unambiguously positive.

JEL Classification: J20, J30, C23

Keywords: linked employer-employee panel data, fixed effects, limited mobility bias

Corresponding author:

Thorsten Schank  
Johannes Gutenberg-Universität Mainz  
Jakob-Welder-Weg 4  
55128 Mainz  
Germany  
E-mail: [schank@uni-mainz.de](mailto:schank@uni-mainz.de)

# 1 Introduction

Positive assortative matching (PAM) implies that high productivity workers and firms match together. This intuitively plausible idea goes back to Roy (1951), Becker (1973), Sattinger (1975), but more recent contributions include Kremer (1993) and Shimer & Smith (2000). The extent to which PAM is actually observed in the labour market sheds light on the mechanisms which determine matching, and has important policy implications, not least because PAM is related to the degree of wage inequality.

Following the publication of Abowd, Kramarz & Margolis (1999) a number of papers have attempted to find evidence for PAM by estimating wage equations with worker and firm fixed effects. However, the majority of the literature has found small or even negative correlations between the worker and firm effects.<sup>1</sup>

There are three possible explanations for this stylised fact. First, there are a number of highly structural models that attempt to model this prediction. In particular, some papers suggest that it is difficult to identify assortative matching from wage data only (de Melo 2008, Eeckhout & Kircher 2011). See also Bagger & Lentz (2008). This prompted Mendes, van den Berg & Lindeboom (2010) to take a more direct approach and estimate plant-level production functions using observable proxies for worker and firm productivities. They find clear evidence of PAM.

Second, it is possible that two-way fixed-effects wage equations are misspecified because they ignore the contribution of additional worker-firm match effects. Woodcock (2008) estimates a wage equation which allows for such match effects, and finds that the estimated correlation between worker and firm effects increases from zero to 0.185.

The third explanation is that there is a limited mobility bias in the estimated cor-

---

<sup>1</sup>See, for example, Goux & Maurin (1999), Abowd, Kramarz, Perez-Duarte & Schmutte (2009), Woodcock (2008), and Gruetter & Lalive (2009).

relation caused by estimation error. This was noted originally by Abowd, Kramarz, Lengeremann & Perez-Duarte (2004) but was developed formally by Andrews, Gill, Schank & Upward (2008). Andrews et al. develop formulae that show that the estimated correlation is biased downwards if there is true PAM. Moreover, this bias is bigger the fewer workers who move between firms in the data, which is why it is labelled limited mobility bias.

In this paper we show empirically that limited mobility bias matters a lot for the estimated correlation between worker and firm effects. Using the employment statistics register of the German Federal Office of Labour we alter the amount of inter-firm mobility by sampling a varying fraction of workers from the population. We show that the estimated correlation between worker and firm contributions to wage equations is negative when inter-firm mobility is small, but the correlation becomes unambiguously positive for larger samples with more inter-firm mobility.

## 2 Methodology and Limited Mobility Bias

Using linked employer-employee panel data, the literature typically estimates

$$y_{it} = \mu + \mathbf{z}_{1it}\boldsymbol{\beta}_1 + \mathbf{z}_{2jt}\boldsymbol{\beta}_2 + \theta_i + \psi_j + \varepsilon_{it}. \quad (1)$$

There are  $i = 1, \dots, N$  workers,  $j = 1, \dots, J$  firms and  $t = 1, \dots, T$  years.  $y_{it}$  is wages;  $\mathbf{z}_{1it}$  is a vector of observable time-varying worker covariates and  $\mathbf{z}_{2jt}$  is a vector of observable time-varying firm covariates.  $\theta_i$  and  $\psi_j$  are time-invariant (scalar) unobserved heterogeneities, potentially correlated with each other, but also with  $\mathbf{z}_{1it}$  and  $\mathbf{z}_{2jt}$ . Workers may move between firms; there are  $M$  movers in total.

It is standard to assume strict exogeneity:

$$\mathbb{E}(\varepsilon_{it} | \mathbf{1}, \mathbf{z}_{1i1}, \dots, \mathbf{z}_{1iT}, \mathbf{z}_{2j1}, \dots, \mathbf{z}_{2jT}, \theta_i, \psi_j) = 0.$$

This implies workers' mobility decisions are independent of  $\varepsilon_{it}$ , but can be a function of the unobservables  $\theta_i$  and  $\psi_j$ . Because the  $\theta_i$  and  $\psi_j$  are correlated with the observed covariates, random effects methods are biased and inconsistent, and so two-way fixed-effects methods are needed to estimate  $\theta_i$  and  $\psi_j$ .<sup>2</sup>

Evidence for PAM comes from seeing whether or not the correlation between the worker and firm components of Equation (1) is positive:

$$\text{Corr}(\theta_i, \psi_j) > 0 \quad \text{or} \quad \text{Corr}(\mathbf{z}_{1it}\boldsymbol{\beta}_1 + \theta_i, \mathbf{z}_{2jt}\boldsymbol{\beta}_2 + \psi_j) > 0. \quad (2)$$

These two correlations each comprise a covariance and two variances, which in turn depend on  $\theta_i$  and  $\psi_j$ . As noted by Krueger & Summers (1988) both  $\text{Var}(\hat{\theta}_i)$  and  $\text{Var}(\hat{\psi}_j)$  are biased upwards; this is because every  $\theta_i$  and every  $\psi_j$  are subject to estimation error. Andrews et al. (2008) show that  $\text{Cov}(\hat{\theta}_i, \hat{\psi}_j)$  is also biased, because of estimation error and because the estimates of  $\theta_i$  and  $\psi_j$  are related by:

$$\hat{\theta}_i - \theta_i = -\bar{\mathbf{z}}_{1i}(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) - \bar{\mathbf{z}}_{2i}(\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_2) - (\bar{\hat{\psi}}_i - \bar{\psi}_i) + \bar{\varepsilon}_i,$$

where “ $\bar{\cdot}$ ” averages a variable over  $t$ , and “ $\hat{\cdot}$ ” denotes an estimate from Equation (1). Conditional on the observed covariates, if a  $\psi_j$  is over-estimated, then, on average, the corresponding  $\theta_i$  is under-estimated, and *vice versa*. Thus, if the true correlation is positive, then the estimated correlation is biased downwards. Further, Andrews et al. (2008) show formally, and by simulations, that the bias can be sizeable, and reduces as the the number of workers who move between firms,  $M$ , increases.

---

<sup>2</sup>Standard estimation methods are not practical when the number of firms is large. We use `a2reg` in Stata 11 (Ouazad 2008), which implements the conjugate gradient algorithm method of Abowd, Creecy & Kramarz (2002).

### 3 Data and Empirical Strategy

The data come from the employment statistics register of the German Federal Office of Labour (*Beschäftigtenstatistik*), which covers all workers or trainees registered by the social insurance system (Bender, Haas & Klose 2000). Each observation has a unique establishment identification number.<sup>3</sup> We select all workers in the employment register who were employed on June 30th each year to create a simple annual unbalanced panel, 1998-2007. To keep sample sizes manageable, we use the two most populous states in Western Germany (Bavaria and North Rhine Westphalia) and the most populous in Eastern Germany (Saxony).<sup>4</sup>

The original sample sizes are approximately 88m (Bavaria) 122m (North-Rhine Westphalia) and 15m (Saxony). From these samples we select full-time workers aged 16-65 who work in the private sector and who have non-missing values for  $y_{it}$ ,  $\mathbf{z}_{1it}$  and  $\mathbf{z}_{2it}$ .<sup>5</sup> We then keep only those observations which belong to the largest interconnected group, where a group contains all the workers who have ever worked for any of the establishments in that group, as well as all the establishments at which any of those workers were employed. A second (unconnected) group is defined only if no establishment in the first group has ever employed any workers in the second and no establishment in the second group has ever employed any workers in the first.<sup>6</sup> Now the sample sizes are approximately 46m, 62m and 7m respectively.

The dependent variable is daily gross wages, which are censored at the social security contribution ceiling.<sup>7</sup> These censored observations will also attenuate the estimated

---

<sup>3</sup>Typically the literature refer to firms when discussing PAM; our data, in fact, comprise establishments.

<sup>4</sup>Splitting the sample means that we lose inter-establishment mobility which occurs between states. However, the degree of inter-state mobility is extremely low. Between 2006 and 2007 the proportion of workers remaining in the same state is 98.5% (Bavaria), 98.5% (North-Rhine Westphalia) and 97.6% (Saxony).

<sup>5</sup>In our estimates of Equation (1)  $\mathbf{z}_{1it}$  comprises tenure and a set of occupation dummies,  $\mathbf{z}_{2it}$  comprises log establishment size. We also include year dummies which capture the effects of time, age and experience.

<sup>6</sup>The largest group accounts for 97.8% of remaining observations.

<sup>7</sup>The proportion of observations in our samples which are censored in 2007 are 11.0% (Bavaria),

correlation between worker and establishment effects towards zero.

The results of Andrews et al. (2008) suggest that  $\text{Corr}(\hat{\theta}_i, \hat{\psi}_j)$  should be increasing and concave in the number of movers per establishment  $M/J$ , asymptoting towards the true correlation. When sampling real data, one can increase  $M$  by increasing either the proportion of workers sampled, the proportion of establishments sampled, or the number of time periods sampled. However, there may be genuine effects of PAM in the data that confound the relationship between the bias and the number of movers. For example, increasing the proportion of workers sampled changes the size distribution of the sample of establishments, and the true correlation between worker and establishment effects may vary with establishment size.

To get a clean experiment that allows us to increase the number of movers, but keep the sample of establishments constant, we:

1. Take a 10% random sample of workers, and define  $p$  as the proportion of workers sampled ( $p = 0.1$ );
2. Record the identities of all establishments which employ those workers;
3. Holding this sample of establishments constant, increase  $p$  to 0.2, 0.3, 0.5 and 1.

We do not take a random sample of establishments and vary  $p$ , because we would lose all inter-establishment mobility to and from establishments outside the sample.

The  $p = 0.1$  sample results in  $J = 65,032$  (Bavaria), 84,564 (North-Rhine Westphalia) and 19,877 (Saxony). Table 1 summarises the sample sizes and worker movements observed in those samples of establishments when we increase  $p$ . Thus, for example, in Bavaria we observe an average of 29.6 worker movements per establishment over the period 1998–2007 when all workers are sampled.

---

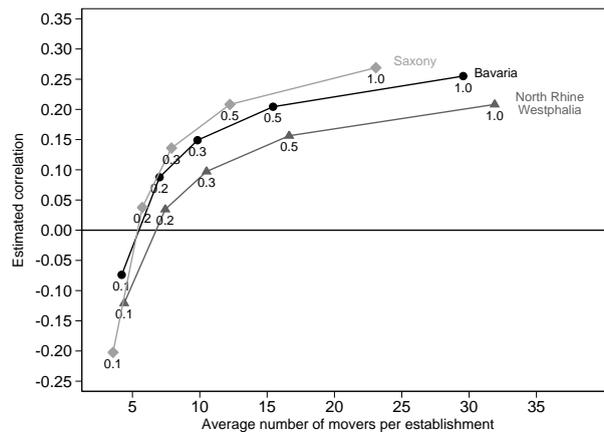
9.5% (North-Rhine Westphalia) and 4.7% (Saxony).

$p$	<i>Bavaria</i> $J = 65,032$		<i>North-Rhine Westphalia</i> $J = 84,564$		<i>Saxony</i> $J = 19,877$	
	$N^*$	$M/J$	$N^*$	$M/J$	$N^*$	$M/J$
0.1	1,779,562	4.2	2,309,319	4.4	436,766	3.6
0.2	3,393,479	7.0	4,409,560	7.4	820,059	5.7
0.3	5,003,038	9.8	6,519,154	10.5	1,205,597	7.9
0.5	8,214,938	15.4	10,735,633	16.6	1,977,795	12.2
1.0	16,278,473	29.6	21,270,334	31.9	3,904,445	23.1

**Table 1:** Increasing the proportion of workers sampled in a fixed sample of establishments increases the number of worker movements per establishment.

## 4 Results

Our basic results are reported in Figure 1. Each data point represents a single regression and resulting  $\text{Corr}(\hat{\theta}_i, \hat{\psi}_j)$ . The proportion of workers sampled is also indicated. The correlation increases strongly with  $p$ , and the pattern matches very closely the simulated results presented in Andrews et al. (2008, Figure 1). The effect of increasing  $p$  is very consistent across all three states, even though one of those states (Saxony) is in Eastern Germany which has not yet completed the transformation process.



**Figure 1:** Increasing the number of movers per establishment in a fixed sample of establishments increases  $\text{Corr}(\hat{\theta}_i, \hat{\psi}_i)$ .

This demonstrates that there is a positive correlation between worker and estab-

lishment effects in German data.<sup>8</sup> These correlations are only slightly smaller than those estimated by Mendes et al. (2010) for Portugal using data on productivity. Our results also explain why many studies do not find such a correlation, because any given dataset could have been sampled anywhere along the  $M/J$ -axis. Indeed, our results also explain why some studies, with very few movers, estimate negative correlations.

The inclusion of observable characteristics in the correlation (see Equation 2) does not change our conclusion because the observable components are not subject to limited mobility bias. When we estimate  $\text{Corr}(\mathbf{z}_{1it}\boldsymbol{\beta}_1 + \theta_i, \mathbf{z}_{2jt}\boldsymbol{\beta}_2 + \psi_j)$ , the correlation increases by only 0.04 (Bavaria), 0.03 (North-Rhine Westphalia) and 0.02 (Saxony) when  $p = 1.0$ .

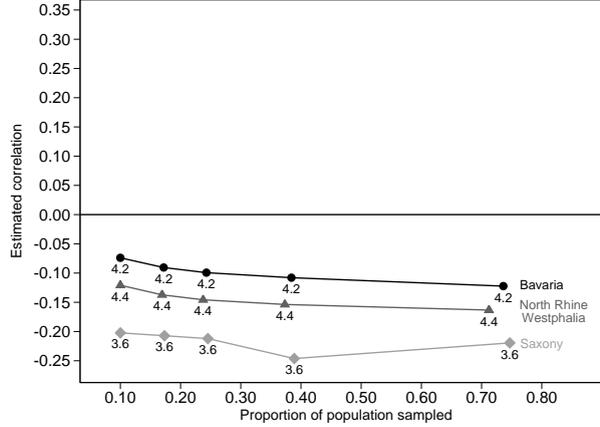
The increase in  $\text{Corr}(\hat{\theta}_i, \hat{\psi}_j)$  is not simply a result of increasing the sample size (worker-years). To show this, we repeat the experiment of increasing  $p$ , but now we only keep additional workers if they do not join or leave their establishment during the sample period. This ensures that the number of movers per establishment is held fixed. The result is shown in Figure 2, which plots the resulting proportion of workers sampled (which is necessarily less than  $p$ ) against the estimated correlation. The (fixed) number of movers per establishment is also indicated at each point. Because  $M/J$  is held constant there is no increase at all in  $\text{Corr}(\hat{\theta}_i, \hat{\psi}_j)$ .

## 5 Conclusion

The existing empirical literature has generally failed to find a positive correlation between worker and firm components of wage equations, a result often seen as evidence against PAM. We show that limited mobility bias can have a large effect on the estimated correlation if the data include only a small number of worker movements

---

<sup>8</sup>The estimate for  $p = 1.0$  is still likely to be a lower bound because wages are top-censored.



**Figure 2:** Increasing the number of observations per establishment, but keeping the number of movers constant does not increase  $\text{Corr}(\hat{\theta}_i, \hat{\psi}_i)$ .

per firm. In our data, when the number of movers per establishment is small ( $< 5$ ) the estimated correlation is consistently negative. When the number of movers per establishment is large ( $> 25$ ), the estimated correlation is consistently positive and in the range  $0.2 - 0.3$ . This strongly suggests that, for Germany, the true correlation is positive.

## Acknowledgements

The authors thank the IAB (Institut für Arbeitsmarkt und Berufsforschung, Nürnberg), in particular Lutz Bellmann, for funding this research whilst Richard Upward was working at the Institute. The views expressed in this paper are solely those of the authors and are not those of the IAB. The comments of Daniel Hamermesh and Martin Biewen and other participants at presentations in the economics departments at Mainz, Manchester, Nottingham and Sheffield are gratefully acknowledged.

## References

- Abowd, J., Creecy, R. & Kramarz, F. (2002), Computing person and firm effects using linked longitudinal employer-employee data, Technical Paper 2002-06, U.S. Census Bureau.
- Abowd, J., Kramarz, F., Lengermann, P. & Perez-Duarte, S. (2004), Are good workers employed by good firms? A test of a simple assortative matching model for France and the United States. Mimeo, CREST.
- Abowd, J., Kramarz, F. & Margolis, D. (1999), ‘High wage workers and high wage firms’, *Econometrica* **67**, 251–333.
- Abowd, J., Kramarz, F., Perez-Duarte, S. & Schmutte, I. (2009), A formal test of assortative matching in the labor market. NBER Working Paper 15546.
- Andrews, M., Gill, L., Schank, T. & Upward, R. (2008), ‘High wage workers and low wage firms: negative assortative matching or limited mobility bias?’, *Journal of the Royal Statistical Society, Series A* **171**, 673–97.
- Bagger, J. & Lentz, R. (2008), An empirical model of wage dispersion with sorting. mimeo.
- Becker, G. (1973), ‘A theory of marriage: part I’, *Journal of Political Economy* **81**, 813–846.
- Bender, S., Haas, A. & Klose, C. (2000), ‘The IAB employment subsample 1975-1999’, *Schmollers Jahrbuch* **120**, 649–662.
- de Melo, R. (2008), Sorting in the labour market: theory and measurement. Job market paper.
- Eeckhout, J. & Kircher, P. (2011), ‘Identifying sorting—in theory’, *Review of Economic Studies* **78**, 872–906.

- Goux, D. & Maurin, E. (1999), ‘Persistence of interindustry wage differentials: a reexamination using matched worker-firm panel data’, *Journal of Labor Economics* **17**, 492–533.
- Gruetter, M. & Lalive, R. (2009), ‘The importance of firms in wage determination’, *Labour Economics* **16**, 149–160.
- Kremer, M. (1993), ‘The O-ring theory of economic development’, *Quarterly Journal of Economics* **108**, 551–575.
- Krueger, A. & Summers, L. (1988), ‘Efficiency wages and the inter-industry wage structure’, *Econometrica* **56**, 259–93.
- Mendes, R., van den Berg, G. & Lindeboom, M. (2010), ‘An empirical assessment of assortative matching in the labor market’, *Labour Economics* **17**, 919–929.
- Ouazad, A. (2008), A2REG: Stata module to estimate models with two fixed effects. Statistical software components, Boston College Department of Economics.
- Roy, A. (1951), ‘Some thoughts on the distribution of earnings’, *Oxford Economic Papers* **3**, 135–46.
- Sattinger, M. (1975), ‘Comparative advantage and the distribution of earnings and abilities’, *Econometrica* **43**, 455–468.
- Shimer, R. & Smith, L. (2000), ‘Assortative matching and search’, *Econometrica* **68**, 343–669.
- Woodcock, S. (2008), Match effects. Mimeo, Simon Fraser University.