2018

Augmenting the Human Capital Earnings Equation with Measures of Where People Work

Erling Barth
Institute for Social Research and National Bureau of Economic Research, erling.barth@samfunnsforskning.no

James C. Davis
US Bureau of the Census and Boston Federal Statistical Research Data Center, james.c.davis@census.gov

Richard B. Freeman
Harvard University and National Bureau of Economic Research, freeman@nber.org

Augmenting the Human Capital Earnings Equation with Measures of Where People Work

Erling Barth, *Institute for Social Research and National Bureau of Economic Research*

James Davis, *US Bureau of the Census and Boston Federal Statistical Research Data Center*

Richard B. Freeman, *Harvard University and National Bureau of Economic Research*

We augment standard log earnings equations for workers in US manufacturing with variables reflecting measured and unmeasured attributes of their employer. Using panel employee-establishment data, we find that establishment-level employment, education of co-workers, capital equipment per worker, and firm-level R&D intensity affects earnings substantially. Unobserved characteristics of employers captured by employer fixed effects also contribute to the variance of log earnings, although less than unobserved characteristics of individuals captured by individual fixed effects. The observed and unobserved measures of employers mediate the effects of individual characteristics on earnings and increase earnings inequality through sorting of workers among establishments.

We have benefited from support from the Labor and Worklife Program at Harvard University, the National Bureau of Economic Research, and the Norwegian Research Council (projects 202647, 236786, and 179552 [Erling Barth]). Thanks to Thomas Lemieux for very useful comments. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the US Census Bureau. All results have been reviewed to ensure that no confiden-

[*Journal of Labor Economics*, 2018, vol. 36, no. S1]
 0734-306X/2018/36S1-0010$10.00
Submitted July 27, 2016; Accepted June 1, 2017

I. Introduction

Standard earnings equations relate log earnings of individuals to their measured human capital or demographic attributes.[1] These equations account for a sizable proportion of the variation in individual earnings but still leave a sufficiently large residual among workers with the same measured characteristics to challenge "the law of one price" in the US labor market. Exemplifying the dispersion of earnings for workers with similar skills, Devroye and Freeman (2001) found that variance of log earnings among US workers within narrow bands of adult literacy test scores exceeded the variance of earnings among all workers in the United Kingdom, the Netherlands, and Sweden. Much of the increased earnings inequality in the United States from the 1970s to the 2000s, moreover, has taken the form of increased inequality among workers with the same measured human capital or demographic attributes.

What underlies the level and change in the residual variance from log earnings equations? One identifiable factor is the firm or establishment that employs the worker. The simplest market-clearing models postulate either a single wage or a narrow band of wages associated with costs of information and mobility among jobs for workers with comparable skills, but the evidence often shows that employers pay sizable differences for workers with the same measured attributes.[2] Commensurately, the same worker often earns substantially more or less working for different employers in closely aligned periods of time.

II. ANOVA of Establishment Earnings

To see the extent to which earnings of individuals depend on where they work in the United States in recent years, we analyzed the log earnings of individual workers in the Longitudinal Employer-Household Dynamics (LEHD) Employment History Files for the nine states with LEHD data from 1992 through 2007.[3] We decomposed the total variance of log earnings into a part attributable to differences in earnings among establishments and a part due to differences in earnings of workers within establishments. To keep as many observations as possible on each individual for panel data anal-

---

tial information is disclosed. Contact the corresponding author, Richard B. Freeman, at freeman@nber.org. Information concerning access to the data used in this paper is available as supplementary material online.

[1] We use "attributes" and "characteristics" interchangeably in this paper.

[2] Earlier studies include Davis and Haltiwanger (1991), Groshen (1991), Abowd, Kramarz, and Margolis (1999), Lane, Salmon, and Spletzer (2007), Gruetter and LaLive (2009), Nordström Skans, Edin, and Holmlund (2009), Card, Devicienti, and Maida (2014), Barth et al. (2016), and Song et al. (2015) as well as the contributions in Lazear and Shaw (2009).

[3] See Sec. III.A below for details.

Table 1
Variance Decomposition of Log Earnings, All Sectors, 1992 and 2007

	Variance Log Earnings		Share of Variance between Establishments, 2007	Change in Variance, 1992–2007	Share of Growth between Establishments, 1992–2007
	1992	2007			
All	.510	.601	.48	.091	.66
Mining, utilities, and transport	.434	.457	.40	.022	.39
Business services	.612	.713	.56	.101	.86
Communication	.502	.634	.40	.132	.53
Retail, wholesale, and restaurants	.508	.551	.48	.044	.80
Finance, insurance, and real estate	.531	.660	.39	.129	.65
Private services	.427	.482	.49	.054	.90
Health, education, and social services	.495	.508	.27	.013	−.15
Manufacturing	.398	.490	.45	.092	.57

NOTE.—Numbers are calculated from yearly regressions of log annualized sum of quarterly earnings for all jobs in the second quarter of the year on establishment dummies. Data are from the Longitudinal Employer-Household Dynamics. Establishment is the SEINUNIT.

ysis, we included all jobs observed in the second quarter of each year in the LEHD. We estimated establishment effects by regressing the log earnings of individuals on establishment dummies separately for each year and then used the variance of the estimated coefficients on establishment dummies to measure the variance of log earnings due to the between-establishment effect. The remaining variance reflects earnings differences within establishments and interactions between individuals and establishments. Since the ANOVA does not adjust earnings for the measured attributes of individuals within establishments or for establishment differences in average worker attributes or observable establishment attributes, the calculations are a descriptive representation of the raw earnings data.

Table 1 displays the results of the ANOVA for 1992–2007 for the whole US economy and for eight large one-digit sectors. The columns give the total variance of individual earnings in 1992 and 2007, the share of the variance attributable to variance among establishments, the change in variance over time, and the share of the change attributable to increased variance of earnings among establishments. The first line shows that in the economy as a whole 48% of the variance of log earnings among workers comes from variation among establishments and that 66% of the 0.091 increased variance of earnings is due to the increase in variance among establishments.[4]

[4] These estimates are nearly identical to those in Barth et al. (2016) based on analysis of full-year main jobs, which found that 49% of variance of log earnings was

The remaining lines of table 1 show differences in the level and change in the variance of log earnings among sectors. Variance is highest in business services and in finance, insurance, and real estate and is lowest in manufacturing; in mining, utilities, and transportation; and in private services. The share of the total variance associated with establishments is largest in business services and private services and is lowest in health, education, and social services. Manufacturing, on which this paper focuses, has lower variance of earnings than the economy as a whole, is close to the economy-wide share of variance associated with earnings differences among establishments, and has a lower establishment share of the 1992–2007 increase in variance.

What employer attributes determine whether an employer pays above- or below-average market wages? How much do the individual attributes in standard log earnings equations and unobserved individual attributes associated with earnings affect the contribution of the employer effects on the variances shown in table 1? What economic forces compress earnings across employers? What forces increase divergence? Do earnings differentials associated with worker characteristics differ by employer enough to contribute to the overall dispersion of earnings?

We examine these questions for manufacturing. We focus on manufacturing because of the quality of data on that sector: the Annual Survey of Manufacturers (ASM) provides information on manufacturing establishments annually that is unavailable for other sectors and detailed evidence on investment in capital and other inputs that are likely to affect labor productivity across establishments, potentially leading to heterogeneity of pay through some kind of rent-sharing mechanism.[5] For our analysis, we combined individual earnings from the LEHD with data on worker attributes from the decennial census and the Current Population Survey (CPS), data on establishment attributes from the Census of Manufacturing, and data on firm attributes from the Longitudinal Business Database (LBD) and the Survey of Industrial Research and Development (SIRD).[6]

---

associated with variance of earnings in 2007 and that 67% of the increase in labor earnings inequality from 1992 to 2007 was due to the increased variance of earnings among establishments.

[5] The literature on rent sharing relates individual earnings to establishment/firm productivity or profitability; see, e.g., Hellerstein, Neumark, and Troske (1999), Margolis and Salvanes (2001), Dunne et al. (2004), Faggio, Salvanes, and van Reenen (2010), Dobbelaere and Mairesse (2010), Mortensen, Christensen, and Bagger (2010), and Card, Heining, and Kline (2013).

[6] These data measure firm employment, establishment employment, capital per worker, percentage of output exported overseas, and R&D per employee. We use data on the individuals in each establishment to estimate the average characteristics of the establishment workforce: years of schooling, age, gender, and race.

III. Methodology

The traditional human capital earnings equation (Mincer 1974) relates earnings w_{it} of individual i in period t to observable measures/indicators of personal skill and other individual characteristics that ideally reflect productivity but that may also reflect employer attitudes or perceptions resulting from prejudicial or statistical discrimination:

$$\log w_{it} = \beta_0 + \gamma_t + x_{it}\beta + u_{it}, \tag{1}$$

where γ_t is a period effect and x_{it} is a vector that includes years of schooling and individual attributes such as age, gender, and race. The equation does not include attributes of the establishment or firm, although they can be added to the equation to reflect compensating differential or other factors related to the full compensation of workers that are not captured by the earnings measure.

Our augmented earnings equation adds the measured and unmeasured characteristics of an individual's establishment/firm to equation (1):

$$\ln w_{ijt} = \beta_0 + \gamma_t + x_{it}\beta + z_{j(it)t}d + \psi_{ij(it)} + e_{ijt}, \tag{2}$$

where $j(it)$ is an index of the workplace that employs individual i at time t and $\psi_{ij(it)}$ is a unique job fixed effect for every individual and workplace pair. For clarity of exposition, we omit the (it) index and write only the index j to indicate the workplace that employs individual i at time t in the following. The t subscript on z_{jt} allows observed employer characteristics to vary over time at a workplace, which could potentially affect the earnings of workers. Our analysis assumes that an employer characteristic affects the earnings of all workers similarly. With a panel of workers and employers, the d coefficients for the establishment characteristics are estimable using within-job variation in employer characteristics—for example, if z relates to employment ("larger establishments pay more"), the effect of employment on earnings can be estimated for workers in the same job when the establishment changes employment.

Having multiple observations on a person along with employer identifiers in longitudinal data allows us to decompose the job effect into an individual fixed effect via a dummy variable for each worker, an establishment fixed effect via a dummy variable for each employer, and a match component orthogonal to the individual and establishment fixed effects per the Abowd-Kramarz-Margolis (AKM) decomposition (Abowd et al. 1999): $\psi_{ij} = \alpha_i + \phi_j + \xi_{ij}$. Defining $\alpha_i = X_iB + a_i$ and $\phi_j = Z_jD + \varphi_j$, where X and Z are covariates for each individual and establishment, we identify the B and D parameter vectors by assuming that the residual of the individual fixed effect is orthogonal to individual fixed characteristics and that the residual of the establishment fixed effect is orthogonal to establishment fixed characteristics. However, the components of both fixed effects can correlate with time-varying characteristics and each other. Our final augmented equation is

$$\begin{aligned}\ln w_{ijt} &= \beta_0 + \gamma_t + x_{it}\beta + X_i B + a_i + z_{jt}d + Z_j D + \varphi_j + \xi_{ij} + e_{ijt} \\ &= \beta_0 + \gamma_t + \omega_{it} + \Omega_{jt} + \xi_{ij} + e_{ijt},\end{aligned} \tag{3}$$

where ω_{it} is the individual component of the earning and Ω_{jt} is the establishment component, both of which contain observable and unobservable parts.

Comparing equations (1) and (3), if personal skills and attributes are the sole factors included in the estimation, the coefficients of equation (1) estimate the gross return to those skills/attributes inclusive of possible gains from access to different employers, whereas the coefficients of equation (3) measure the net return exclusive of the earnings characteristics of employers. Alternatively, to the extent that the covariates in equation (1) are correlated with the equation (2) variables, the estimated coefficients of (1) can be viewed as biased estimates of the net effects of skills/attributes in equation (2).

A. Matched LEHD, Establishment, and Firm Data

As noted, our dependent variable is the earnings of individual workers in the LEHD Employment History Files for the nine states with LEHD data from 1992 through 2007.[7] We link the LEHD to the quinquennial Census of Manufacturers (CoM) for the economic census years 1992, 1997, 2002, and 2007 and to the ASM in intermediate years, using the LEHD Business Register Bridge that links data at the firm level. LEHD establishments are linked by firm, detailed industry, and county to CoM/ASM establishments. For the vast majority of observations—single-unit firms and plants of a firm located in a different county than other plants of the firm in the same industry—the mapping from LEHD to CoM/ASM establishments is unique within detailed industry and county. But for plants of a firm in the same industry and county, the link is not one to one. For these establishments, we aggregate plant characteristics to the firm-industry-county level and link these measures to their workers.

The COM/ASM data provide production-related data on manufacturing establishments, which we add to the files on employees: number of workers

[7] The LEHD data provide annualized quarterly earnings from the unemployment insurance benefit programs, linked to the Quarterly Census of Employment and Wages. We only use observations that include positive earnings in the second quarter of the year. Abowd, Creecy, and Kramarz (2002) describe the construction of the LEHD data. The nine states are California, Colorado, Idaho, Illinois, Maryland, North Carolina, Oregon, Washington, and Wisconsin. They cover approximately half of US employment. Comparisons with data for states that cover different time periods show that the nine-state sample is reasonably representative (Barth et al. 2016).

at establishments and establishment capital equipment and building stock as constructed by Foster, Grim, and Haltiwanger (2016) with perpetual inventory methods. We measure firm employment from the LBD and whether a firm reports R&D expenditures and the amount from the SIRD. Table A1 gives summary statistics for our key variables.

We obtain measures of the years of schooling, occupation, age, race, and gender of workers in the LEHD by linking workers to their characteristics in the 1990 and 2000 decennial census long form and March CPS files for 1986–1997. The Census Center for Administrative Records staff matched these data using the protected identification key (PIK) identifier, which is the person identifier in the Census, the CPS, and the LEHD. Beginning with 2000, decennial files have very high PIK match rates, of 90%–93% (Mulrow et al. 2011; Rastogi and O'Hara 2012). However, the 1990 PIK is more limited due to the vintage of address files.[8] Matching Census/CPS data to the LEHD Employment History Files provides us with data on years of schooling and other worker attributes for 20.5% of employees in the LEHD data.[9] In the matched sample, we require that a person is observed at least four times.

Table A1 shows that the matching process produces a sample that is higher in earnings and worker attributes that are positively associated with earnings, such as age and being white, and a sample that is also higher in firm and establishment attributes that are positively associated with earnings, such as number of employees and capital per employee.

IV. Variance Decomposition in Manufacturing

Table 2 gives a variance decomposition for the subset of manufacturing workers for whom we match observations in the LEHD and CoM to decennial census or CPS files. This is the sample on which the rest of our analysis focuses. The increase in variance in the subsample falls short of the increase in the full LEHD—a variance of 0.272 compared with the table 1 figure of 0.398 in 1992 and a variance of 0.330 compared with 0.490 in 2007, producing a smaller increase in variance over time. A major reason is that the matched sample loses many small establishments where earnings are relatively low. The proportion of the variance in log earnings attributable to between-establishment differences is as a result lower as well. The 43% contribution of increased earnings between establishments in the matched sample is smaller than the 57% in the full LEHD. The matched

[8] Individual name and address files are highly sensitive and are not generally distributed in the US Census Bureau with the data files. Our versions of 1990 decennial files did not have original name and address data and had to be reconstructed with other data sets. As a result, the PIK matches favor less mobile adult heads of household.

[9] We first matched to the 2000 Census, then matched missing cases to the 1990 Census, and finally matched missing cases to the CPS data.

Table 2
Variance Decomposition of Log Earnings in the Matched Longitudinal Employer-Household Dynamics (LEHD) Manufacturing Panel, 1992 and 2007

	Variance, 1992	Share	Variance, 2007	Share	Change, 1992–2007	Share of Change
Log earnings	.272	1	.330	1	.058	1.00
Between establishments	.125	.46	.150	.45	.025	.43
Between firms	.113	.42	.140	.42	.027	.47
Between establishments within firm	.012	.04	.011	.03	−.001	−.03
Within establishments	.146	.54	.180	.55	.033	.57

NOTE.—Numbers are calculated from a regression of log earnings on time dummies and establishment dummies. The matched sample includes LEHD data matched to the Census of Manufacturers with valid observations of capital (from the Annual Survey of Manufacturers/Census of Manufacturers tfp files; see Foster et al. 2016) and to education data from the decennial censuses and the Current Population Survey; each individual is observed at least four times (for details, see Sec. III.A). All jobs included are observed in the second quarter of the year. Slight differences between the "Between establishments" numbers and the sum of "Between firms" and "Between establishments within firms" numbers are due to rounding errors.

sample understates the contribution of establishments to the variation in earnings.

As establishments belong to firms that include other establishments, all of which may be covered by firm-wide human resource and compensation policies, we decomposed the between-establishment contribution to the variance of earnings into an effect associated with firms and an effect associated with establishments within firms. We did this in two stages: first by regressing log of earnings on dummy variables for establishments and then by regressing estimated establishment fixed effects on dummy variables for firms. The proportion of the variance attributed to firms reflects the overall pay practices of firms, while the remaining proportion reflects pay differences among establishments in the same firm.

The table 2 calculations show that consistent with the emphasis of Song et al. (2015) on the importance of the firm in accounting for the increased dispersion in worker earnings over time, the firm component dominates the variation in log earnings among establishments. In our manufacturing data, 90.4% (=0.113/0.125) of the variance in earnings between establishments in 1992 is assigned to firm fixed effects and 93.3% (=0.140/0.150) of the establishment variance in 2007 is similarly assigned to firm fixed effects. Over time, moreover, the variance in establishment earnings for establishments within firms fell, so that 110% (=0.47/0.43) of the increased earnings dispersion associated with establishment was due to increased earnings variance among firms.[10]

[10] Because many small firms have only a single establishment, the calculation that assigns virtually all of the variance of single-establishment firms to the firm could overstate the dominance of firms in establishment effects. To see how much of the

V. Cross-Section Earnings Equations

Column 1 of table 3 records estimated coefficients and standard errors for ordinary least squares regressions of the benchmark cross-section log earnings equation with years of schooling, age, gender, and some interactions to allow for differences in effects among those attributes. In addition, the regression includes 171 geographic area dummies and 16 year dummies so that the coefficients are estimated within year and area. The estimated coefficients are similar to those typically found in the human capital earnings literature: an estimated average return to years of schooling of about 9.4% per year and a concave age profile captured by the negative squared term and gender and race earnings gaps at 30% and 17%, respectively. The R^2 of the equation of 0.45 is larger than the R^2 in earnings functions fit on CPS data,[11] presumably because variation in earnings in the entire economy exceeds that in manufacturing and/or because the administrative LEHD earnings has less measurement error than self-reported earnings in the CPS.

Column 2 adds a set of workplace variables to reflect place of employment: four-digit NAICS industry dummies, the log number of employees of the firm, the log number of employees in the establishment, and establishment age and its square.[12] The estimates show significant firm and establishment effects and a concave earning-establishment age profile. Adding the firm and establishment characteristics raises the R^2 to 0.505 and thus explains 10% of the residual variance of earnings for demographically similar persons. The firm and establishment variables shrink the positive coefficients on years of schooling and age and the negative coefficients on gender and being nonwhite, indicating that some of the impact of those factors comes through sorting of workers among establishments and industries within manufacturing.

Column 3 adds variables relating to the attributes of the establishment's workforce: mean years of schooling, mean age, share female, and share nonwhite; capital structures per worker and capital equipment per worker; the export share of establishment revenues; and the R&D investment of the firm to which the establishment belongs. The most striking result is the high estimated coefficient on the years of schooling of all workers. The estimated 0.069 coefficient on the mean years of schooling in the workers' establishment compared with the 0.074 coefficient on the workers' own education

table 2 result is due to single-unit firms, we eliminated them from the data set and decomposed the variances of earnings among multiunit establishments. Table A2 shows that among multiestablishment firms, 83% (=0.094/0.113) of the variation in establishment fixed effects is associated with the firm fixed effect, which supports the conclusion in the text.

[11] Estimating a similar regression with CPS data for the whole workforce gave an R^2 of 0.35.

[12] Dickens and Katz (1987) examine industry wage differentials. Brown and Medoff (1989) study employer size-wage effects.

Table 3
Estimated Regression Coefficients and Standard Errors for Augmented Earnings Equations, Including Firm and Establishment Characteristics for Manufacturing, 1992–2007

	Model 1: No Establishment Characteristics (1)	Model 2: Establishment Characteristic I (2)	Model 3: Establishment Characteristic II (3)	Model 4: Plus Firm Fixed Effects (4)	Model 5: Plus Establishment Fixed Effects (5)
Years of schooling	.0943***	.0820***	.0737***	.0739***	.0739***
	(.0001)	(.0001)	(.0001)	(.0001)	(.0001)
Age	.0132***	.0118***	.0116***	.0116***	.0115***
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
Age^2	−.0005***	−.0005***	−.0005***	−.0005***	−.0005***
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
Female	−.3022***	−.2897***	−.2691***	−.2671***	.2669***
	(.0006)	(.0006)	(.0005)	(.0005)	(.0005)
Nonwhite	−.1707***	−.1539***	−.1405***	−.1397***	−.1400***
	(.0005)	(.0005)	(.0005)	(.0005)	(.0005)
Female × Age	−.0050***	−.0045***	−.0045***	−.0045***	−.0044***
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
Female × Age^2	.0000***	.0000***	.0000***	.0000***	.0000***
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
Female × Nonwhite	.0309***	.0318***	.0384***	.0371***	.0373***
	(.0009)	(.0009)	(.0008)	(.0008)	(.0000)
Establishment and firm characteristics:					
Log firm employment		.0295***	.0173***	.0071***	.0012***
		(.0001)	(.0001)	(.0005)	(.0003)
Log establishment employment		.0266***	.0270***	.0258***	.0232***
		(.0002)	(.0002)	(.0003)	(.0006)
Establishment age		.0032***	.0045***	.0032***	
		(.0001)	(.0001)	(.0001)	

Establishment age^2		−.0001***	−.0001***	−.0001***	−.0000***
		(.0000)	(.0000)	(.0000)	(.0000)
Mean years of schooling			.0690***	.0483***	.0249***
			(.0003)	(.0004)	(.0007)
Mean age			.0027***	.0017***	−.0034***
			(.0001)	(.0001)	(.0001)
Share female			−.3176***	−.2548***	−.1084***
			(.0016)	(.0029)	(.0056)
Share nonwhite			−.0185***	.0029	.0256***
			(.0017)	(.0028)	(.0046)
Export share			.0115***	.0062***	.0068***
			(.0004)	(.0005)	(.0006)
Log capital structures/employee			.0074***	.0033***	.0018***
			(.0002)	(.0003)	(.0004)
Log capital equipment/employee			.0462***	.0211***	.0075***
			(.0003)	(.0004)	(.0005)
Firm R&D/employee			.7260***	.1228***	.1627***
			(.0104)	(.0131)	(.0129)
Industry effects		Y	Y	Y	
Firm effects				Y	
Establishment effects					Y
Region (171) effects	Y	Y	Y		
Year (16) effects	Y	Y	Y	Y	Y
r^2 adjusted	.452	.505	.526	.566	.577
N	5.13E+06	5.13E+06	5.13E+06	5.13E+06	5.13E+06

*** $p < .001$.

suggests that it is almost as good to work in an establishment with more educated workers as it is to have more education. The estimates also show that workers earn more in establishments with older workers and less in establishments with a larger proportion of female or nonwhite workers. More capital equipment per worker raises earnings more than more capital structures per worker (a coefficient difference of 0.046 vs. 0.007), and earnings are higher in establishments with a high export share. Finally, earnings rise with R&D intensity of a firm: workers in firms with 1 standard deviation higher R&D intensity average 2% more earnings.

Column 4 gives the regression results with dummy variables for firms added to the equation, while column 5 gives results with dummy variables for establishments replacing those for firms. With establishment fixed effects in the model, it is no longer possible to identify separately the linear effect of establishment age and the time dummies, so we have removed the linear term for establishment age. Since the effect of linear age is now absorbed by the time dummies, none of the remaining estimators are affected. Addition of the firm fixed effects substantially reduces the estimated impact of the number of employees at the firm, indicating that short-run changes in firm employment have little effect on earnings but only reduce the coefficient on number of employees at the establishment modestly. The column 5 estimates with dummy variables for establishment also markedly reduce the coefficient for firm employment but leave a substantial effect of establishment employment on earnings. With establishment fixed effects in the equation, the positive effect of establishment employment suggests that an establishment operates along a rising supply curve of labor for short-term increases in employment, which suggests some monopsony power in the labor market (Manning 2005).

Addition of the firm and establishment dummies naturally shrinks the estimated effect of firm and establishment variables on earnings. The column 4 firm fixed effects regression eliminates the negative relationship between the share of nonwhite employees. Working in an establishment with a large nonwhite share is associated with low earnings, but short-run changes in the nonwhite share do not affect establishment earnings much. The column 4 fixed effects regression also greatly weakens the relationship between R&D and earnings, reducing the estimated coefficient by more than 80%. While R&D firms pay more than firms that do less R&D, changes in R&D activity within a firm have little effect on earnings.

The column 5 regression, which includes establishment fixed effects, further shrinks the coefficients of most of the establishment workforce characteristics compared with those in column 3. The estimated 0.0690 effect of the mean years of schooling on earnings in column 3 drops to 0.0249 in column 5, while the estimated −0.3176 for being female in column 3 drops to −0.1084 in column 5. While measurement error usually accounts for some of the lower coefficient on variables in longitudinal analysis compared with cross-section analysis (Freeman 1984), the pooling of observations to create average characteristics is likely to diminish measurement error so that huge drops in the effects of these characteristics are likely to at least in part reflect economic behavior as firms adjust earnings to changing characteristics gradually over time.

VI. Panel Earnings Equations

The longitudinal structure of the LEHD allows us to estimate the effects of employer characteristics on earnings for the same individual in two ways: (1) by comparing workers who remain in the same job while management changes characteristics of the establishment or does nothing to offset changes due to factors outside management control, such as workers retiring or quitting for another job without replacing the leaver with someone similar, and (2) by comparing workers who quit an employer with one set of characteristics to join an employer with other characteristics. Outside of recession years, the bulk of the labor mobility comes from worker decisions to move to a new employer willing to hire them. In recessions, mobility depends more on the layoff decisions of firms, with the number of layoffs increasing to approach or exceed the number of quits.[13] While our data lack information on whether a worker left a job by quitting or by layoff, recession years are less frequent than nonrecession years in our data, which suggests that the bulk of the worker changes reflect quits rather than layoffs.[14]

Table 4 presents estimates of the effect of employer attributes on the earnings of the same worker when those attributes change. Column 1 shows the results of adding individual fixed effects to the basic log earnings regression from column 3 in table 3 for all workers in the matched sample. The coefficients on some employer variables decline with the addition of the worker fixed effects: the estimated coefficient for average years of schooling of workers in an establishment falls by 59% (from 0.0737 to 0.0299), suggesting that much of the large coworker schooling effect is due to positive sorting of workers by unmeasured individual characteristics into establishments with more educated workers. The coefficient on the equipment stock of capital per employee drops more massively by 70% (from 0.0462 to 0.0140), suggesting positive sorting of unmeasured individual characteristics into establishments with more equipment capital. And the coefficient on R&D drops by 83% (from 0.762 to 0.1290), suggesting that most of the cross-section

[13] For nonrecession years, the number of quits divided by the number of layoffs exceeds 1.0 by 30%–50%. In recession years, the number of layoffs exceeds quits. See chart 7 in BLS (2015).

[14] We did not probe possible differences between job changes from establishments having large drops in employment, where layoffs are potentially important, and job changes from establishments with stable or growing employment, where the locus would likely be voluntary shifts to better outside opportunities.

Table 4
Estimated Regression Coefficients and Standard Errors for Firm and Establishment Characteristics: Individual Fixed Effects and Job Fixed Effects Models

	Model 1: Individual Fixed Effects, All Workers (1)	Model 2: Individual-Job Fixed Effects, Stayers (2)	Model 3: Individual Fixed Effects, Movers (3)
Log firm employment	.0161***	.0056***	.0162***
	(.0001)	(.0002)	−.0003
Log establishment employment	.0305***	.0214***	.0180***
	(.0002)	(.0003)	−.0005
Establishment age	.0037***		−.0008***
	(.0001)		(.0002)
Establishment age square	−.0001***	−.0000***	.0000
	(.0000)	(.0000)	(.0000)
Mean years of schooling	.0299***	.0048***	.0322***
	(.0003)	(.0004)	−.0008
Mean age	.0023***	−.0129***	.0023***
	(.0002)	(.0001)	−.0002
Share female	−.1573***	−.0036	−.1765***
	(.0020)	(.0033)	−.0044
Share nonwhite	.0250***	.1025***	.0230***
	(.0019)	(.0029)	−.0042
Export share	.0000	−.0021***	.0020*
	(.0003)	(.0003)	−.0009
Log capital structures/employee	.0059***	.0002	.0031***
	(.0002)	(.0003)	−.0004
Log capital equipment/employee	.0140***	.0027***	.0189***
	(.0002)	(.0003)	−.0006
Firm R&D expenses/employee	.1290***	.0860***	.2075***
	(.0060)	(.0057)	−.0188
r^2 adjusted	.873	.913	.827
N	5.13E+06	5.13E+06	7.31E+05

Note.—All models include year dummies, age^2, the interaction between gender and age, and the interaction between gender and age^2. The first and third models include individual fixed effects, and the second model includes job (i.e., match: the unique combination of individual and establishment) fixed effects. Individual-specific variables that do not vary over time, such as years of education, are absorbed by the individual fixed effects.
* $p < .05$.
*** $p < .001$.

R&D effect is due to a positive matching between R&D firms and unmeasured individual characteristics.

The next two columns unpack the fixed effects model into its two parts. Column 2 estimates the effect of employer characteristics on the earnings of workers who stay in the same establishment. This specification controls for what we call "job-individual fixed effects"—the unique combination of an individual and the establishment, which encompasses both the establishment fixed effects and the individual fixed effect from the AKM decomposition, in addition to a potential match-specific fixed effect. As in table 3, the linear effect of establishment age is no longer separately identified from the time effects and is thus omitted from the model. Column 3 estimates the impact of employer factors on the earnings of workers who changed employers, which identifies the effects of establishment characteristics through changes in the employer and thus does not control for establishment fixed effects or match-specific effects.[15]

For most establishment characteristics, the column 3 estimated effects of worker-initiated changes have a much greater impact on earnings than do the column 2 estimated effects of employer-initiated changes. Moving to a firm that has greater employment gives an earnings increase of 0.0162, while working in a firm that increases employment changes gives a 0.0056 boost to earnings—about one-third as large. Moving from an establishment with more years of schooling increases earnings by 0.0322, compared with an increase in earnings of 0.0048 when a worker's current establishment increases its years of schooling. Moving to an establishment with older workers raises the earnings of the mover, while staying in an establishment with a rising age of the workforce reduces the worker's earnings. The effect of R&D on earnings is more than twice as large for movers than for stayers (0.2075 vs. 0.0860). But not all characteristics have a larger effect for movers than for stayers. An increase in establishment employment has a modestly larger effect for persons who stay with an establishment than for those who move, and similarly for the share of nonwhites.

Mechanically, the differences between the column 2 stayers-based estimates and the column 3 movers-based estimates reflect the fact that the stayers analysis controls for unobserved establishment fixed effects and thus removes correlations between those effects and the earnings, while the movers model does not do this. But the differences also reflect economic behavior. A worker who chooses to change employers will likely require a larger increase in pay to cover the costs of mobility than one who stays at a job. An establishment that changes characteristics will likely adjust operations slowly and alter pay less in the short run compared with employers whose characteristics differ over longer periods and whose pay structures reflect long-term differences in the mode of operating.[16]

Earnings equations with individual fixed effects cannot identify the relation between stable individual characteristics and earnings: those effects are

[15] For this analysis we examine every job-to-job move in the data, retaining only the observations before and after the move, and include individual fixed effects in the regression.

[16] Measurement error will also bias downward the estimates based on changes, for the basic reason that a given error will have a proportionately larger impact on the small variation in year-to-year changes at the same workplace than on the larger differences between the employer the worker joins and the employer the worker leaves.

Table 5
Regression of Estimated Individual Fixed Effects on Years of Schooling and Demographic Individual Characteristics from Three Models of Log Earnings

	Model 1: Fixed Effects from Model with Individual Characteristics Only (1)	Model 2: Fixed Effects from Model with Individual and Establishment Characteristics (2)	Model 3: Fixed Effects from Model with Individual and Establishment Characteristics and Establishment Fixed Effects (3)
Years of schooling	.1076***	.0917***	.0841***
	(.0001)	(.0001)	(.0001)
Dummy variable for female	−.3553***	−.3326***	−.3129***
	(.0004)	(.0004)	(.0004)
Dummy variable for nonwhite	−.1001***	−.0882***	−.1119***
	(.0005)	(.0005)	(.0004)
Age	.0136***	.0128***	.0112***
	(.0000)	(.0000)	(.0000)
Gender × Nonwhite interaction	.0504***	.0500***	.0429***
	(.0009)	(.0008)	(.0007)
r^2 adjusted	.441	.422	.416
Variance of the estimated unobserved individual effects	.149	.127	.112

NOTE.—The dependent variable is the individual fixed effects from models including time-varying covariates. Number of observations is 5.13E+06 for all columns.
*** $p < .001$.

absorbed in the individual dummy variables. But it is possible to learn something about how years of schooling and demographic factors such as gender, race, or age affect the individual fixed effects by regressing the estimated fixed effect for individuals on those characteristics. Say we have 10 workers with two defining characteristics, years of schooling and gender. The fixed effects earnings equation would produce estimated coefficients for each of the 10 workers that could be regressed on the workers' schooling and gender to capture their relation to the fixed effects. Columns 1–3 of table 5 give the results of such an analysis in three regression models. Model 1 uses estimated individual fixed effects from a regression without employer characteristics.[17] Model 2 uses estimated individual fixed effects from a regression with observable employer characteristics. Model 3 uses estimated individual fixed effects from a stayers' regression that includes establishment and match-specific fixed effects (eq. [3]). The estimated relations between the individual effects that are positively related to the characteristics of employers decline across the columns as we add increasing information

[17] The difference is that in the fixed effects specification, the unobserved individual fixed effects are allowed to be correlated with all of the included time-varying covariates.

Table 6
Variance Decomposition of the Full Augmented Earnings Equation Model

Determinants of Earnings	Variance Decomposition
Log earnings	.299
Individual components	.188
Observed individual	.078
Unobserved individual	.112
2 × Cov(observed individual, unobserved individual)	−.002
Within-match residual	.020
Establishment components	.043
Observed establishment	.024
Unobserved establishment	.020
Unobserved firm	.016
Unobserved establishment within firm	.005
2 × Cov(observed establishment, unobserved establishment)	−.001
2 × Cov(individual components, establishment components)	.038
2 × Cov(observed individual, observed establishment)	.022
2 × Cov(observed individual, unobserved establishment)	.012
2 × Cov(unobserved individual, observed establishment)	.008
2 × Cov(unobserved individual, unobserved establishment)	−.006
Match component	.010

NOTE.—Calculations used eq. (3) to structure decomposition. Number of observations is 5.13E+06. Data are for the manufacturing matched sample, as described in the text. Year dummies are not included in the calculations. Some numbers do not add up due to rounding errors.

on where the employee works. The returns to years of schooling drops from 0.1076 for the model 1 specification that has no controls for employee characteristics to 0.0841 for the model 3 specification that controls for observed and unobserved establishment effects. The coefficient on female falls by 12%, and the coefficient on age falls by 18%.

The bottom line in table 5 (“Variance of the estimated unobserved individual effects”) shows how the addition of establishment characteristics reduces the contribution of the fixed effects for individuals to the variation of earnings among workers. In model 1, the individual fixed effect variance is 0.149, or 51% of the total variance. In model 2, which includes measured establishment characteristics, the individual fixed effect variance falls to 0.127, or 43% of the total variance. In model 3 with observed and unobserved establishment characteristics, the variance of the individual effect is 0.112, or about 38% of the total variance in earnings. Put differently, establishment factors account for 25% ((0.149 − 0.112)/0.149) of the variance of estimated individual effects.

VII. A Full Decomposition

Table 6 summarizes our findings with a full decomposition of log earnings in the augmented earnings equation. The standard individual characteristics of years of schooling, age, gender, and race account for 26% of the total variation in earnings; unobserved individual effects account for 37% of the variation; observed establishment characteristics account for 8% of the variance; unobserved establishment effects account for 7% of the variance; and the match component accounts for 3% of the variance. The covariance between the individual and establishment components of the earnings equation adds 13% of the variance. The remaining variation arises from the transitory within-match residual comprising 7% of the total variation and to small negative covariance terms between the observed and unobserved parts of the individual and establishment components, respectively.[18] The most important factors relate to individuals, but employer factors and their relation to individual factors are significant and substantive.

VIII. The Sorting of Workers between Establishments

The interaction between individual characteristics and establishment characteristics suggests that sorting of workers with given characteristics among workplaces with different characteristics affects inequality at large. Positive assortative matching of workers high in measured or unmeasured skills/wages to high-wage establishments raises the inequality of earnings. By contrast, assortative matching of workers with workers of similar measured skill does not create "extra inequality" but points to the complementarity of skills of similar workers in the production process and allocation of labor.

Assortative matching also affects the interpretation of estimated coefficients on particular variables. When workers positively sort by education into higher-paying establishments, the traditional log earnings equation that excludes establishment factors captures two effects in its estimated coefficient on years of schooling: the return of higher skills to earnings within an establishment and the differential access that schooling gives workers to obtaining jobs in higher-paying establishments. Addition of dummy variables for establishments limits the effect of years of schooling to its effect within an establishment. Given that sorting of workers between establishments affects the dispersion of pay and the returns to individual characteristics, we analyze next the ways in which workers and firms match up.

Table 7 gives the correlation coefficients for sorting by key earnings determinants. The largest correlations show considerable sorting of workers with workers like themselves: correlations of educated workers with educated workers (0.477), of older workers with older workers (0.333), of females with females (0.349), and of nonwhites with nonwhites (0.471). But

[18] Our model assumes that the fixed individual and establishment/firm effects remain constant throughout the sample period. Experiments with estimation on subperiods show that in fact the variance of both the individual fixed effects and the establishment fixed effects appear to rise during the sample period. The period over which to treat individual and establishment/firm fixed effects as fixed raises statistical and modeling issues and merits further analysis.

Table 7
Correlation Coefficients between Individual and Establishment/Firm Characteristics

	Years of Schooling	Age	Female	Nonwhite
Log firm employment	.200	.060	−.002	−.059
Log establishment employment	−.028	.149	−.014	−.028
Establishment age	.214	.034	.015	−.042
Export share	.103	.044	.005	−.054
Log structures capital/employee	.185	.069	−.059	−.087
Log equipment capital/employee	.117	.071	−.100	−.075
Firm R&D/employee	.216	.015	.000	.000
Mean years of schooling	.477	.031	−.030	−.128
Mean age	.110	.333	−.060	−.036
Share female	−.033	−.041	.349	.105
Share nonwhite	−.146	.005	.080	.471
Establishment observables as a group, weighted by effect on earnings	.258	−.022	−.084	.030
Establishment fixed effect	.112	.069	−.061	−.066

NOTE.—Coefficients are tabulated from the matched data file for manufacturing workers, as described in the text.

other characteristics of employers are sufficiently correlated with worker characteristics to suggest sorting of workers among establishments beyond homophily. Educated workers work in large firms and in R&D-intensive firms, in establishments with high capital per worker and high export shares. These patterns make it likely that some of the education earnings premium comes through the greater likelihood that educated workers find jobs in employers with other earning-enhancing characteristics. Older workers are also associated with establishments with high-earning characteristics, although the correlations are much smaller. By contrast, women work in establishments with lower capital intensity, and nonwhite workers are largely employed in establishments with low-earning characteristics.

The bottom two lines of the table shows the correlation between a composite measure of the establishment contribution to earnings through observed variables plus industry and region, weighted by their estimated effect on earnings, and through establishment fixed effects. Both the establishment observables and the fixed effects are highly correlated with years of schooling, making schooling potentially the most important dimension of worker sorting among establishments.

Figure 1 summarizes the relations between the characteristics of workers and those of the establishments where they work via the correlations between indices of the observed characteristics as a group, weighted by their respective coefficients in the earnings equation, and the fixed effects associated for workers as well as for establishments. The largest correlation is between the individual observables weighted by their contribution to earnings and establishment observables weighted by their contribution to earnings

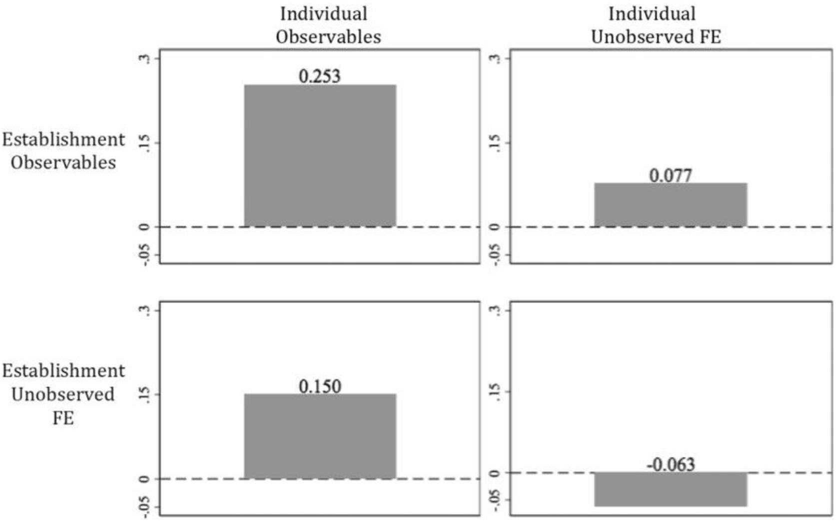

FIG. 1.—Correlation coefficients between individual observables and unobservables with establishment observables and unobservables. Coefficients are calculated from the matched data file for manufacturing workers, as described in the text. FE = fixed effect. A color version of this figure is available online.

(0.253), followed by the correlation between the individual observables and the establishment unobserved fixed effect. By contrast, the fixed effect of individuals is weakly positively correlated with the establishment observables, while the individual unobservables and unobserved establishment fixed effects are negatively correlated—a result consistent with Abowd et al. (2014). As Andrews et al. (2008) notes, a negative correlation between two unobserved components of earnings could result from sampling and measurement errors,[19] so the safest conclusion from these correlations is that sorting of workers occurs largely on observable characteristics.

IX. Mobility across Employers

The impact of employer characteristics on the earnings of workers with similar measured characteristics and fixed effects and the table 7 and figure 1 correlations direct attention to the potential role of worker mobility among employers in determining pay. To what extent does mobility from job to job raise pay? How often do workers who start their careers in establishments with low-earning characteristics move to firms with better observable and unobservable characteristics over time? And conversely, how much downward firm mobility is there among workers who begin their careers at firms with high-earning characteristics?

[19] Lise, Meghir, and Robin (2013) provide further discussion of these issues.

To examine the transitions of workers among establishments with different establishment components of earnings, we formed a transition matrix for workers in our data in both 1992 and 2007. We attached to every worker the total establishment contribution to earnings, defined as the sum of the contribution to earnings of the time-varying establishment characteristics, such as firm size and R&D spending; the fixed observables, such as industry and region; and the unobserved establishment effects. With an establishment contribution for each worker in 1992 and 2007, the natural measure of each worker's mobility is the change in the establishment component of earnings of their employer in those years.

Table 8 summarizes the transition pattern by quintiles of the distributions, ordered from low-paying firms in quintile 1 to high-paying firms in quintile 5. The rows in the table show the distribution of workers by the 1992 quintile distribution of their employer into the 2007 quintile distribution of their employer. While the largest probabilities are for workers to remain in the same quintile over time, there is evidence of upward movement among establishments. Workers in the low quintiles have larger shares going up in the distribution than workers in the top quintiles have shares falling in the distribution. Among workers in the third quintile, 38% move to a higher quintile, whereas only 21% move down and 40% remain in the same quintile. New workers come into the distribution of firms at the lower end and change jobs over time to produce a lifetime move up the distribution. Measured by productivity of establishments rather than earnings, Haltiwanger, Hyatt, and McEntarfer (2018, in this issue) find a similar pattern in productivity, with younger workers in particular moving from less productive to more productive firms over time.

Finally, we characterize the sorting of workers with workers between establishments by Kremer and Maskin's (1996) index of segregation, $\rho = \mathrm{cov}(\omega\, \underline{\omega})/V(\omega)$, where $\underline{\omega}$ is the average individual component of the establishment and $V(\omega)$ is the variance of the individual components of the

Table 8
Transitions of Workers among Establishments Ordered by Establishment Contribution to Earnings (Observable Characteristics Weighted by Their Earnings Coefficients Plus Fixed Effect) by Quintile of the Distribution, 1992–2007

	Quintile (2007)						Change in Share	
1992	1	2	3	4	5	All	Up	Down
Quintile 1	.564	.258	.097	.056	.024	1.000	.436	
Quintile 2	.172	.401	.287	.108	.032	1.000	.427	.172
Quintile 3	.078	.136	.403	.329	.054	1.000	.383	.214
Quintile 4	.042	.062	.113	.451	.331	1.000	.331	.217
Quintile 5	.017	.024	.033	.136	.790	1.000		.210

NOTE.—Data are calculated on the balanced panel only. Quintiles of the distribution of the establishment effect include both unobserved and observed components of the establishment contribution.

standard earnings equation. If workers segregate completely between establishments according to their individual earnings components, $\rho = 1$, whereas if they randomly allocate between establishments, $\rho = 0$. The index of segregation for observable characteristics in our data is 0.24, while the index for unobservable characteristics is 0.17. This supports the implication of the correlations that sorting of workers according to observed individual characteristics, such as years of schooling, age, gender, and race, is considerably stronger than segregation according to unobserved attributes.

We characterize the sorting of workers with establishments by the equivalent Kremer-Maskin (1996) index, $\rho_\Omega = \text{cov}(\omega, \Omega)/V(\omega)$, where Ω is the earnings components of establishment and divide the decomposition into its within-establishment (V^w) and between-establishment (V^b) parts by the identities

$$V^w(\log w) = V(\omega)(1 - \rho) + V(\xi) + V(e), \tag{4a}$$

$$V^b(\log w) = V(\omega)(\rho + 2\rho_\Omega) + V(\Omega). \tag{4b}$$

In our data, $\rho = 0.247$ and $\rho_\Omega = 0.100$. The within-establishment component is 59% of the variance in earnings, of which 82% arises from the observed and unobserved individual component, 6% arises from the match component, and 12% arises from the residual. The between component contributes 41% of the variance in earnings, of which 36% is due to worker-worker sorting, 30% is due to worker-establishment sorting, and 34% is due to variance of the establishment effect. That most of the within-establishment variation in earnings is associated with variables related to individuals and most of the between-establishment variation in earnings is associated with variables related to establishments and sorting of workers among establishments suggests that the simple within and between decomposition offers powerful insight into the role of supply and demand factors in earnings determination.

X. Conclusion

Augmenting the earnings equation with measured characteristics of employers and unobserved earnings-related fixed effects for establishment or firm adds substantially to the variance in log earnings explicable by an earnings equation. The regressions identify observable employer-side factors—capital equipment per worker, R&D investments, export performance, the level of schooling of an establishment's workforce, the number of employees, and, up to a point, the age of the establishment—associated with higher worker earnings. While the sizable cross-section effects of measured employer characteristics diminish in longitudinal data with the inclusion of firm and establishment fixed effects, those fixed effects are another manifestation of the importance of where a person works to what they earn.

The evidence that addition of establishment- or firm-related variables reduces the estimated coefficient on the key human capital variable, years of schooling, by about one-fifth directs attention to the role of schooling in giving workers access to higher-paying workplaces. The near comparability of the estimated effect of average establishment schooling and of the individual's years of schooling (absent firm or establishment fixed effects) further suggests that some of the gains from human capital investments spill over to other workers,[20] while the drop in the estimated coefficient on average years of schooling with the addition of establishment fixed effects reflects the strong positive relation between average years of schooling and establishment fixed effects.

The evidence that the estimated coefficients on gender, race, and age diminish when we introduce individual and establishment fixed effects provides further support for the notion that the sorting of workers among employers is important in earnings differentials. Youth, females, and nonwhites are more likely to be found in low-paying establishments, and the coefficients for the establishment-specific average of the demographic characteristics change signs or turns insignificant once we control for job fixed effects. While the dynamics of worker mobility has workers moving to enterprises with higher observable and fixed effects earnings components over time, assortative matching tends to magnify the effect on inequality beyond what would be the direct impact of earnings differences across establishments.

Taken together, the dual findings that where a person works affects their earnings and that sorting of workers among employers accounts for some of the differentials in earnings associated with years of schooling and demographic characteristics raises new questions for analysis: Why does having more educated coworkers affect individual earnings so much? To what extent do costs of mobility account for the greater impact of employee-instituted than employer-instituted changes in the payoff from working with other inputs? How important are explicit human resource and compensation policies in positioning employers in the distribution of earnings? And do their decisions equilibrate the marginal payoffs to paying above- or below-market average levels of pay? Finally, given manufacturing's modest and declining share of employment, our analysis suggests the value of estimating augmented earnings functions in other industries to see whether the role of employers and sorting of workers found here generalizes to the labor market writ large.

[20] The relation between the average years of schooling at an establishment and individual pay is mindful of the observed positive relation between the average level of education at a regional or city level and earnings, which has been interpreted in terms of human capital externalities (see, e.g., Moretti 2004).

Appendix

Table A1
Summary Statistics: Mean and Standard Deviation (SD) of Variables for the Full Sample of Persons in the Longitudinal Employer-Household Dynamics (LEHD) Manufacturing Data Set and in the Matched Sample That Includes Measures of Years of Schooling

	Full Sample		Matched Sample	
	Mean	SD	Mean	SD
Years of schooling			12.72	2.311
Age	42.58	10.183	43.41	9.978
Female	.300	.458	.294	.456
Nonwhite	.302	.459	.237	.425
Log firm employment	8.234	2.403	8.288	2.268
Log establishment employment	6.300	1.572	6.317	1.512
Log capital structures/employee	3.250	1.333	3.297	1.296
Log capital equipment/employee	3.918	1.049	3.953	1.035
Export share of establishment	.626	.484	.638	.481
Firm R&D/employee	.009	.021	.009	.020
Log earnings	6.649	.554	6.682	.543
Observations (in millions)	23.4		5.1	

NOTE.—The full sample is tabulated for all workers in the LEHD; the matched sample is tabulated for workers reporting years of schooling in match with the Census or the Current Population Survey, as described in the text.

Table A2
Variance Decomposition of Earnings among Workers, in All Firms and in Multiestablishment Firms, Matched Panel, 1992–2007

	All	Share	MUs	Share
Log earnings	.293	1	.279	1
Between establishments	.119	.41	.113	.41
Between firms	.106	.36	.094	.34
Between establishments		.04		.06
Within firm	.013		.018	
Within establishments	.174	.59	.166	.59

NOTE.—Firms in the "All" establishment sample include the establishment effects for single-unit firms, whereas the "MUs" sample includes only multiestablishment firms (defined as multiunit firms within manufacturing only). Numbers are calculated from a regression of log earnings on time dummies and establishment dummies. The total variance is calculated after subtracting variance due to the time dummies. Firm effects are estimated from regression of establishment fixed effects on firm dummies. Multiunit firms are defined as multiunit firms within manufacturing only.

References

Abowd, John M., Robert H. Creecy, and Francis Kramarz. 2002. Computing person and firm effects using linked longitudinal employer-employee data. Longitudinal Employer-Household Dynamics Technical Paper no. 2002-06, Center for Economic Studies, US Census Bureau, Suitland, MD.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. High wage workers and high wage firms. *Econometrica* 67, no. 22 (March): 251–334.

Abowd, John M., Francis Kramarz, Sébastien Pérez-Duarte, and Ian M. Schmutte. 2014. Sorting between and within industries: A testable model of assortative matching. NBER Working Paper no. 20472, National Bureau of Economic Research, Cambridge, MA.

Andrews, J., L. Gill, T. Schank, and R. Upward. 2008. High wage workers and low wages: Negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A* 171, no. 3:673–97.

Barth, Erling, Alex Bryson, James C. Davis, and Richard B. Freeman. 2016. It's where you work: Increases in the dispersion of earnings across establishments and individuals in the United States. Special issue, *Journal of Labor Economics* 34, no. 2, suppl. (April): S67–S97.

BLS (Bureau of Labor Statistics). 2015. Job openings and labor turnover survey highlights. http://www.bls.gov/jlt/jlt_labstatgraphs_oct2015.pdf. Washington, DC: BLS.

Brown, Charles, and James Medoff. 1989. The employer size-wage effect. *Journal of Political Economy* 97, no. 5:1027–59.

Card, David, Francesco Devicienti, and Agata Maida. 2014. Rent-sharing, holdup, and wages: Evidence from matched panel data. *Review of Economic Studies* 81, no. 1:84–111.

Card, David, Jörg Heining, and Patrick Kline. 2013. Workplace heterogeneity and the rise of West German wage inequality. *Quarterly Journal of Economics* 128, no. 3:967–1015.

Davis, Steve J., and John Haltiwanger. 1991. Wage dispersion between and within U.S. manufacturing plants, 1963–1986. In *Brookings papers on economic activity, microeconomics*, ed. Martin Baily and Clifford Winston, 115–200. Washington, DC: Brookings Institution.

Devroye, Daniel, and Richard B. Freeman. 2001. Does inequality in skills explain inequality in earnings across advanced countries? NBER Working Paper no. 8140, National Bureau of Economic Research, Cambridge, MA.

Dickens, William T., and Lawrence F. Katz. 1987. Inter-industry wage differences and industry characteristics. In *Unemployment and the structure of labor markets*, ed. Kevin Lang and Jonathan Leonard. Oxford: Basil Blackwell.

Dobbelaere, Sabien, and Jacques Mairesse. 2010. Micro-evidence on rent sharing from different perspectives. NBER Working Paper no. 16220, National Bureau of Economic Research, Cambridge, MA.

Dunne, Timothy, Lucia Foster, John Haltiwanger, and Kenneth R. Troske. 2004. Wage and productivity dispersion in U.S. manufacturing: The role of computer investment. *Journal of Labor Economics* 22, no. 2: 397–430.

Faggio, Giulia, Kjell G. Salvanes, and John van Reenen. 2010. The evolution of inequality in productivity and wages: Panel data evidence. *Industrial and Corporate Change* 19, no. 6 (December): 1919–51.
Foster, Lucia, Cheryl Grim, and John Haltiwanger. 2016. Reallocation in the Great Recession: cleansing or not? *Journal of Labor Economics* 34, no. 1, suppl. (January): S293–S331.
Freeman, Richard Barry. 1984. Longitudinal analyses of the effects of trade unions. *Journal of Labor Economics* 2, no. 1:1–26.
Groschen, Erica. 1991. Sources of intra-industry wage dispersion: How much do employers matter? *Quarterly Journal of Economics* 106, no. 3: 869–84.
Gruetter, Max, and Rafael Lalive. 2009. The importance of firms in wage determination. *Labour Economics* 16, no. 2:149–60.
Haltiwanger, John, Henry Hyatt, and Erika McEntarfer. 2018. Who moves up the job ladder? *Journal of Labor Economics* 36, no. S1 (January): S301–S336.
Hellerstein, Judith K., David Neumark, and Kenneth R. Troske. 1999. Wages, productivity, and worker characteristics: Evidence from establishment-level production functions and wage equations. *Journal of Labor Economics* 178, no. 3:409–46.
Kremer, Michael, and Eric Maskin. 1996. Wage inequality and segregation by skill. NBER Working Paper no. 5718, National Bureau of Economic Research, Cambridge, MA.
Lane, Julia I., Laurie A. Salmon, and James R. Spletzer. 2007. Establishment wage differentials. *Monthly Labor Review* 130, no. 4:3–17.
Lazear, Edward, and Kathryn Shaw, eds. 2009. *The structure of wages: An international comparison*. Chicago: University of Chicago Press.
Lise, Jeremy, Costas Meghir, and Jean-Marc Robin. 2013. Mismatch, sorting and wage dynamics. NBER Working Paper no. 18719, National Bureau of Economic Research, Cambridge, MA.
Manning, Alan. 2005. *Monopsony in motion: Imperfect competition in labor markets*. Princeton, NJ: Princeton University Press.
Margolis, David N., and Kjell G. Salvanes. 2001. Do firms really share rents with their workers? IZA Discussion Paper no. 330, Institute for the Study of Labor, Bonn.
Mincer, Jacob. 1974. *Schooling, experience and earnings*. New York: National Bureau of Economic Research.
Moretti, Enrico. 2004. Human capital externalities in cities. In *Cities and geography*, vol. 4 of *Handbook of regional and urban economics*, ed. J. V. Henderson and J. F. Thisse, 2243–91. Amsterdam: Elsevier.
Mortensen, Dale T., Bent Jesper Christensen, and Jesper Bagger. 2010. Wage and productivity dispersion: Labor quality or rent sharing? 2010 Meeting Papers no. 758, Society for Economic Dynamics.

Mulrow, Edward, Ali Mushtaq, Santanu Pramanik, and Angela Fontes. 2011. Assessment of the U.S. Census Bureau's Person Identification Validation System. NORC at the University of Chicago. http://www.norc.org/PDFs/May%202011%20Personal%20Validation%20and%20Entity%20Resolution%20Conference/PVS%20Assessment%20Report%20FINAL%20JULY%202011.pdf

Nordström Skans, Oskar, Per-Anders Edin, and Bertil Holmlund. 2009. Wage dispersion within and between plants: Sweden 1985–2000. In *The structure of wages: An international comparison*, ed. Edward P. Lazear and Kathryn L. Shaw. Chicago: University of Chicago Press.

Rastogi, Sonya, and Amy O'Hara. 2012. 2010 Census match study. US Census Bureau Report. https://www.census.gov/2010census/pdf/2010_Census_Match_Study_Report.pdf.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. 2015. Firming up inequality. NBER Working Paper no. 21199, National Bureau of Economic Research, Cambridge, MA (revised June 2015).