Does Input Quality Drive Measured Differences in Firm Productivity?*

Jeremy T. Fox
University of Chicago
NBER

Valérie Smeets
Aarhus School of Business
University of Aarhus

February 2009

Abstract

Firms in the same industry can differ in measured productivity by multiples of 3. Griliches (1957) suggests one explanation: the quality of inputs differs across firms. We add labor-market-history variables such as experience and firm and industry tenure, as well as general-human-capital measures such as schooling and sex. We also use the wage bill. We show adding human-capital variables and the wage bill explains only a small portion of productivity dispersion: accounting for input quality decreases the ratio of the 90th to 10th productivity quantiles from 3.74 to 3.36 across six Danish manufacturing industries. The productivity-dispersion decrease is roughly of the same order of magnitude as some competitive effects found in the literature, but input-quality measures do not explain most productivity dispersion, despite economically large coefficient estimates.

JEL codes: D24, L23, M11
Keywords: productivity dispersion, total factor productivity, input quality

*Fox thanks the National Science Foundation, the NET Institute, the Olin Foundation, and the Stigler Center for generous funding. Thanks to helpful discussions with Daniel Ackerberg, Allan Collard-Wexler, Ulrich Doraszelski, Tor Eriksson, Amil Petrin, Chad Syverson, Johannes Van Biesebroeck, and Frédéric Warzynski. We appreciate remarks by seminar participants at Aarhus University, NYU Stern and Texas A&M. Thanks to the professional staff at the Center for Corporate Performance for hospitality and for help integrating the KØB and IDA datasets. Our email addresses are fox@uchicago.edu and vas@asb.dk.

1 Introduction

Measured differences in productivity across plants in the same industry are usually large. Bartelsman and Doms (2000) survey the literature and find many instances where the highest productivity firm has more than twice the measured productivity of the lowest productivity firm. Dhrymes (1995) studies American manufacturing and finds that the ratio of total factor productivity (TFP) of plants in the ninth decile to the TFP of plants in the second decile is 2.75. We find that the ratio of the 90th quantile of productivity to the 10th quantile of productivity is 3.74 in six Danish manufacturing industries. For the same inputs, a firm at the 90th quantile of TFP produces 3.74 times the output of a firm at the 10th quantile of TFP.

These huge differences in cross-sectional, measured productivities have spawned a literature investigating why productivity differences are so large. One explanation is simply measurement error in output. However, measured productivity dispersion is similar in developed and developing countries, whereas measurement error might be expected to be larger in developing-country datasets (Bartelsman and Doms, 2000). Also, productivity at the firm or plant level is persistent across time, meaning any measurement error cannot be transient (Baily, Hulten and Campbell, 1992). Further, productivity dispersion decreases with competition, as theory predicts (Olley and Pakes, 1996; Syverson, 2004; Bloom and Van Reenen, 2007). The literature shows measured productivity predicts firm growth and firm exit (Baldwin, 1995), export success (Bernard and Jensen, 1995), and even transfers of plants between conglomerate firms (Maksimovic and Phillips, 2001; Schoar, 2002). Foster, Haltiwanger and Syverson (2008) use physical output instead of sales as the measure of output, and show that technological productivity dispersion is often even higher than revenue productivity dispersion. Further, both types of productivity are correlated with firm outcomes such as growth and exit. The consensus in the literature is that productivity dispersion is a real phenomenon with important consequences for economic efficiency and our understanding of how markets with heterogeneous producers operate.

This paper investigates whether failing to account for input quality drives productivity residuals. Economists since at least Griliches (1957) have argued that productivity dispersion reflects the quality of inputs across firms. Economists working with US manufacturing-plant data typically measure inputs as the dollar value of physical capital and the number of workers at a firm. Sometimes, employees are separated into production and nonproduction workers. Not surprisingly, labor and capital vary in much greater detail. Two types of machines may have different uses and may not be perfect substitutes, and two types of workers may not have the same contributions to firm output.

Input quality seems to us like a fundamentally different explanation for productivity dispersion than some other explanations, such as management competence, economic incentives, business strategy or other difficult-to-measure firm characteristics. Any firm can buy a higher-quality machine or hire an abler worker simply by paying more money for the higher-quality inputs. If input quality is the reason for productivity, then productivity is really an artifact of a measurement problem. Input markets can be used to reallocate "productivity" across firms: higher-quality workers will switch to the firms that pay them the most, for example. Thus, there is no sense that the firm as an organization is playing an important role in productivity dispersion. If some mostly-fixed firm characteristic such as business strategy explains productivity, then input markets will be less effective at reallocating or increasing productivity. Instead, a Schumpeterian process of creative destruction, where high-productivity firms grow more quickly, may be needed to raise the economy's aggregate productivity. Measuring the role of input quality for productivity dispersion is essential precisely because the optimal policy responses differ depending on whether productivity is due to input quality or some fixed firm characteristic.

As physical capital is measured in monetary units rather than the number of machines, the quality of capital is likely much better measured than the quality of workers in a typical production-function regression. Therefore, our contribution is to disaggregate the labor input. We use matched employer-employee panel data from Denmark to precisely measure many characteristics of workers at a firm. We merge individual-level data on all Danish residents with firm-level data on value added and physical capital. We then construct firm-level statistics about worker characteristics.

We present productivity regressions with increasingly detailed input-quality measures. First, we investigate a simple adjustment, as we follow the literature on income inequality and disaggregate the labor inputs into "skilled" (college) and "unskilled" (noncollege) workers. Next, we include two regressions with much more detailed input quality measures. Schooling, sex, total experience and industry tenure proxy for general- or occupation-specific human capital. Tenure at a worker's current firm proxies for firm-specific human capital. Our production function includes a quality-weighting function that transforms firm-level measures of individual worker characteristics into efficiency units of labor. This labor-quality function is embedded in the estimation of an otherwise standard Cobb-Douglas production function. The residual from this production function estimate is a firm's total factor productivity (TFP). We examine whether adjusting for labor-input quality reduces the measured within-industry dispersion in TFP. We use two different functional forms for labor quality. One specification follows Griliches (1957) and multiplies the contribution of different worker characteristics. The second specification follows Welch (1969) and adds the contribution of each labor quality measure.

Most researchers do not have access to detailed worker panel data. Therefore, we also investigate using proxies for labor quality that can be obtained from accounting data, such as the wage bill of the firm. We present two specifications with wages: the total wage bill instead of the number of workers, and the fraction of the wage bill spent for various human-capital bins. We show that the wage bill reduces productivity dispersion as much as our detailed human-capital measures. Including the wage bill is also interesting because human-capital characteristics tend to have low explanatory power in wage regressions. Here we show that the wage bill does a little better at predicting output than our human-capital measures do; the wage bill may be picking up some unobserved (in our data) worker ability.

We present all of our results separately for two different production functions: Cobb-Douglas and the translog. As firms in different industries use different technologies, we present separate results for eight industries. Also, we present a benchmark for any decline in measured productivity dispersion: the decline in dispersion from adding past employment growth and firm age as controls. The previous empirical literature has emphasized that growth and firm age are correlated with productivity. Another benchmark compares our productivity declines from adjusting for input quality to the productivity declines from local product-market competition in Syverson (2004).

Our main empirical finding is rather startling: especially in manufacturing, our detailed input-quality measures, among the best one can hope to obtain, do not reduce measured productivity dispersion much. The small decrease in productivity dispersion from the detailed labor-quality measures occurs across several types of input-quality measures. The result is robust to the functional form for labor quality and the functional form for the production function. Averaging across six manufacturing industries, the ratio of the outputs of the firm in the 90th quantile of TFP to the firm in the 10th quantile of TFP is 3.74. This declines to 3.36 with human-capital and wage-bill controls. This decline in productivity dispersion is small in terms of its economic magnitude.

Our finding of a small decline in productivity dispersion is not because human-capital measures are unimportant in production. Indeed, for each industry we estimate usually economically-large and sometimes statistically-precise coefficients on the human-capital measures. Rather, our finding is that the reason some firms are dramatically more productive than others is not a simple failure to account for input quality. Adjusting for input quality is not the main explanation for productivity dispersion. As stated above, any firm could hire, say, more college-educated workers simply by paying the market wage for workers with that background. Combining our empirical result with some of the findings from the literature mentioned earlier, the explanation for firm productivity appears to reflect some attribute that is hard to buy and sell in an input market. Explanations include managerial competence, business strategy, or some legally protected competitive advantage. Whatever the true explanation, which the literature is slowly converging upon, the attribute that determines productivity seems to be hard to define and perhaps hard to buy and sell in a market. While discussing optimal policy is well beyond the scope of our paper, this does suggest product-market competition, rather than relying only on input markets, may be an essential force in raising aggregate productivity.[1]

Our main result is that labor quality is not important for productivity dispersion. Given that input quality is listed as one of several possible explanations for productivity in many papers, we feel it is essential to put this theoretically very plausible but empirically untrue hypothesis to rest. Taking the scientific method seriously, no explanation for productivity dispersion can ever be definitively confirmed. Only false hypotheses can be ruled out. As progress in identifying the cause of productivity dispersion comes very slowly, it is important to show negative results of this sort in order to discard these false hypotheses that have been emphasized for at least a half century. Otherwise, the literature cannot make empirical progress.

We have found productivity dispersion is unimportant in manufacturing. Manufacturing is by far the most commonly studied sector in the literature, because of data availability and the intuition that the production processes in manufacturing firms are more comparable across firms in the same industry than in service-sector firms. We study two non-manufacturing industries: hotels and restaurants as well as advertising. For these in-

[1] It is puzzling how low-productivity firms can remain in business at all. One explanation is product differentiation: each firm sells a slightly different product and so heterogeneous consumer demand supports a variety of firms.

dustries, we find economically larger declines in productivity dispersion from including labor-quality measures in production functions. This could be because these industries are more human-capital intensive or it could be that the production technologies in these service-sector industries are more heterogeneous and differences in input qualities proxy for this heterogeneity.

1.1 Literature comparison

Several recent papers use both worker data and firm output data, either to compare production and wage regression coefficients (Van Biesebroeck, 2007) or to control for worker ability in wage regressions (Frazer, 2006). We study productivity dispersion and do not compare our estimates to wages. Likewise, Haltiwanger, Lane and Spletzer (2007) regress TFP residuals on worker-quality controls using US unemployment-insurance data. They focus on the coefficients of labor quality rather than whether productivity dispersion can be explained by input quality.

Hellerstein and Neumark (2006) do find a low increase in R^2.[2] The Danish data we use have more human-capital measures than the corresponding US datasets. We are able to include labor-history measures (experience, industry tenure and firm tenure) constructed from 21 years of panel data for all Danish citizens, and to include a few industries from outside of manufacturing.

Denison (1962) and Jorgenson, Gollop and Fraumeni (1987) account for demographic change (age, race, sex, schooling) and labor quality (they weight demographic groups by wage rates) when decomposing aggregate productivity growth.[3] This pioneering work contrasts with newer empirical work using firm- or plant-level data, which usually does not control for worker quality. Our main finding can be interpreted as saying this change in the controls used in the literature is not substantively important, at least for understanding cross-sectional productivity dispersion.

2 Production, input quality and productivity dispersion

2.1 Production functions

Differences in output across firms can be decomposed into differences in measured inputs, differences in residuals and differences in production technologies. Using data from a single industry and assuming a common

[2] The magnitudes of the productivity-dispersion reduction are hard to compare between the two papers because Hellerstein and Neumark include materials as an input (which raises R^2), while because of data availability (value added is the more common measure in European data), our measure of output is value added, which does not include materials.

[3] A related literature studies the dollar value of accumulated human capital in, for example, US states (Mulligan and Sala-I-Martin, 2000).

technology for all firms, the literature typically estimates the Cobb-Douglas production function

$$\log y = \beta_0 + \beta_l \log l + \beta_k \log k + e, \tag{1}$$

where y is value added, l is the number of workers, k is the monetary value of physical capital, and e is the residual. β_l and β_k are the input elasticities of labor and capital. Between two firms with the same inputs l and k, the firm with the higher output y is said to have a higher measured total factor productivity (TFP), which is $\exp(\beta_0 + e)$ above. Our measure of output is a firm's value added, which is just total sales minus materials and other outsourced inputs, such as consulting services.[4] We focus on e, the productivity residual. We call e productivity throughout the paper.[5]

We also report separate results for the translog production function

$$\log y = \beta_0 + \beta_l \log l + \beta_k \log k + \beta_{l,2} (\log l)^2 + \beta_{k,2} (\log k)^2 + \beta_{l,k} (\log l)(\log k) + e, \tag{2}$$

where the second-order terms and the interaction add approximation flexibility (Christensen, Jorgenson and Lau, 1973). While not reported, our conclusions about TFP dispersion are robust to estimating a constant-elasticity-of-substitution (CES) production function.

2.2 Labor quality

2.2.1 College and noncollege workers

There is only limited work on adding input-quality measures to firm- and plant-level production functions. Therefore, there is no consensus in the literature in how to incorporate input-quality measures. One way is to define new inputs. The empirical literature on income inequality often focuses on "skilled" (workers with a college degree) and "unskilled" (all others) workers. We break the number of workers, l, into $l = l_{\text{college}} + l_{\text{noncollege}}$, where, for example, l_{college} is the number of college-educated workers at a particular firm. We then estimate the Cobb-Douglas production function

$$\log y = \beta_0 + \beta_{\text{college}} \log l_{\text{college}} + \beta_{\text{noncollege}} \log l_{\text{noncollege}} + \beta_k \log k + e.$$

[4] Consistent with the literature, our production functions model the relationship between output and primary inputs like labor and physical capital. We do not have data on intermediate decisions, such as the use of a performance-pay scheme for the workforce. These schemes may indeed raise output, but in production-function language they are intermediate decisions that are concentrated out of the production function. The production function gives output conditional on a firm making appropriate intermediate-input decisions. As we will find a large remaining productivity dispersion when adjusting for labor quality, our results will be consistent with a hypothesis that firms who choose good management practices are more productive.

[5] Like almost every other paper on productivity, for data-availability reasons the dependent variable y is measured in monetary units. Therefore, it incorporates an unmodeled pricing decision. Foster, Haltiwanger and Syverson (2008) do have price data and show that dispersion in technological productivity is actually higher than dispersion in the revenue-productivity measures we work with. Katayama, Lu and Tybout (2006) suggest that supply-and-demand analysis may be more appropriate than productivity analysis when a pricing decision affects the dependent variable.

There is a similar generalization of the translog production function to the case of college and noncollege workers,

$$\log y = \beta_0 + \beta_c \log l_c + \beta_n \log l_n + \beta_k \log k + \beta_{c,2} (\log l_c)^2 + \beta_{n,2} (\log l_n)^2 + \beta_{k,2} (\log k)^2 + \beta_{c,k} (\log l_c)(\log k) + \beta_{n,k} (\log l_n)(\log k) + \beta_{c,n} (\log l_c)(\log l_n) + e,$$

where c and n stand for college and noncollege, respectively.[6]

2.2.2 Human capital measures

Almost every firm in our data has at least one worker with a college degree and one without a college degree. But if there was a firm with no college workers, then $l_{\text{college}} = 0$, $\log l_{\text{college}} = -\infty$ and the firm would produce no output. In a Cobb-Douglas specification, all inputs are essential for production. However, the data show that many firms lack, say, a worker with 3–5 years of tenure at the firm. This means defining $l_{\text{firmtenure},3-5}$ as the number of workers with 3–5 years of tenure and including $l_{\text{firmtenure},3-5}$ as a separate input in a Cobb-Douglas production function contradicts the evidence. Many other types of labor can substitute for those with 3–5 years of tenure; it is not an essential input.

We take several approaches to incorporating more-detailed measures of labor quality into the production function. The first approach follows a classic suggestion of Griliches (1957), who in a survey paper put forth mismeasured input quality as a major explanation for productivity dispersion. This approach views the total labor input as the number of workers times labor quality. Each worker is a bundle of measured characteristics. We unbundle workers so that labor quality is a function of the fraction of workers in a firm with each characteristic.[7] In a firm with 100 workers, hiring 1 more woman with a college degree will increase the fraction of workers who are women by 1% and the fraction of workers with college degrees by 1%. Let $x_{\text{female}} = l_{\text{female}}/l$ be the fraction of workers who are women, and $x_{\text{college}} = l_{\text{college}}/l$ the fraction with a college degree. Total labor quality has the multiplicative functional form

$$q_\theta^{\text{mult}}(x) = (1 + \theta_{\text{female}} x_{\text{female}}) \left(1 + \theta_{\text{college}} x_{\text{college}}\right). \tag{3}$$

Here, efficiency units of labor are the relative productivity compared to a male high-school graduate, say. θ_{female} is how much more productive a woman is than man, and θ_{college} is how much more productive a college-educated worker is than a worker who did not attend college. A firm of all men where 100% of its workers

[6] In a Cobb-Douglas production function, college and non-college workers are complementary inputs: production cannot take place without both inputs. However, there is not a formal model of hierarchical or team production, where perhaps college workers supervise noncollege workers. The flexible translog specification may provide a better approximation to a hierarchical production function.

[7] An exception is total labor-market experience, which enters the labor-quality function as a continuous variable: the mean level of experience in the firm. The data appendix discusses some topcoding reasons why some other variables enter as fractions of the workforce. There is nothing about our method that prevents us from choosing continuous or discrete variables, as appropriate.

attended college will have a per-worker quality of $1+\theta_{\text{college}}$.[8]

Labor quality is not additively separable across workers. For example, expanding the specification of $q_\theta(x)$ above produces the interaction term $\theta_{\text{female}} x_{\text{female}} \theta_{\text{college}} x_{\text{college}}$. If the θ's are positive, adding a male college graduate will produce a greater increase in labor quality at a firm with more women. By contrast, Welch (1969) emphasizes a production technology where human-capital attributes are additive. Therefore, our next functional form for labor quality takes an additive functional form

$$q_\theta^{\text{add}}(x) = 1+\theta_{\text{female}} x_{\text{female}} + \theta_{\text{college}} x_{\text{college}}. \tag{4}$$

Our results about productivity dispersion will be relatively consistent across $q_\theta^{\text{mult}}(x)$ and $q_\theta^{\text{add}}(x)$.

Let the total number of workers at a firm be l. The total labor input is then $l \cdot q_\theta(x)$. Substituting this expression for labor in the Cobb-Douglas production function (1) gives the estimating equation

$$\log y = \beta_0 + \beta_l \log(l \cdot q_\theta(x)) + \beta_k \log k + e. \tag{5}$$

The parameters θ in the labor-quality function enter this equation nonlinearly, so estimation is by nonlinear least squares. This requires using a nonlinear-optimization procedure to minimize the least-squares objective function, as there is no closed-form solution for the least-squares estimator of θ. We also estimate a version of the translog production function, (2), with quality-adjusted labor $l \cdot q_\theta(x)$ replacing the total number of workers, l, as in

$$\log y = \beta_0 + \beta_l \log(l \cdot q_\theta(x)) + \beta_k \log k + \beta_{l,2}(\log(l \cdot q_\theta(x)))^2 + \beta_{k,2}(\log k)^2 + \beta_{l,k}(\log(l \cdot q_\theta(x)))(\log k) + e. \tag{6}$$

The same parameters θ appear in multiple places in the production function.

2.2.3 Wage bill as a proxy for labor quality

Another approach to adjusting for labor quality is to use the wage bill as a measure of the quality of the workforce. Wages will reflect marginal products in a competitive labor market.[9] Just as physical capital is measured in terms of monetary units to reflect the quality of the machinery employed, labor can be measured in terms of its expense in order to reflect its quality. Using the wage bill instead of the number of workers thus makes the methods of measuring physical capital and human capital more symmetric.

The wage bill may also be more commonly found in the type of data used in firm- and plant-level productivity studies. The total wage bill may be part of some accounting-based firm-level datasets where data on the characteristics of the workers are not available. If results from using the wage bill as the labor input are similar

[8] A multiplicative labor-quality measure is also used in Hellerstein and Neumark (2006) and Van Biesebroeck (2007).
[9] Even if the labor market is not perfectly competitive, wages are still likely highly correlated with worker ability.

to those using detailed labor characteristics, then it will ease the data-collection burden for those wanting to control for labor quality.

The wage-bill specification is also attractive because the explanatory power of human-capital variables in wage regressions can be low, suggesting unmeasured worker characteristics are also important determinants of labor quality. Further, the wage bill using monthly salaries better weights the contributions of part-time and full-time workers than do measures like the number of workers.

Our specification with the wage bill is

$$\log y = \beta_0 + \beta_l \log w + \beta_k \log k + e,$$

where the wage bill $w = \sum_{i=1}^{l} w_i$ is the total of the monthly salaries paid to all workers. We also estimate a translog production function, with the wage bill w replacing the number of workers l in (2).

Adding the wage bill could introduce an additional endogeneity problem. If more-productive firms pay higher salaries (for any of several reasons, including profit sharing), then $\mathrm{Cov}(e, w) > 0$ and the benefits of productivity will be misattributed to the labor inputs. Because of this additional endogeneity concern, we do not view the wage bill as a complete replacement for the results with human-capital measures.[10]

2.2.4 Combining the wage bill and the human capital measures

We also combine the wage-bill and human-capital variables to attempt to account for input quality in as detailed a manner as possible. We use the wage bill w instead of the number of workers l as our base labor input. Then we construct a labor-quality adjustment that uses, in part, the human-capital measures. Keeping the same human-capital categories as before, we calculate the total of the monthly wages for workers in each bin and then normalize by the total wage bill of the firm. For example, $\tilde{w}_{\text{female}} = (w)^{-1} \sum_{i=1}^{l_{\text{female}}} w_{i,\text{female}}$ is the fraction of the firm's wage bill that is paid to women. This is a similar measure to x_{female} above, as it represents the fraction of firm labor inputs coming from women. The difference with x_{female} is that the base unit for counting labor inputs is the total of the monthly wages, rather than the number of workers. We then adapt the Griliches (1957) multiplicative-quality-adjustment term, (5), to give

$$q_{\theta}^{\text{mult,wage}}(\tilde{w}) = \left(1 + \theta_{\text{female}} \tilde{w}_{\text{female}}\right)\left(1 + \theta_{\text{college}} \tilde{w}_{\text{college}}\right), \tag{7}$$

where $\tilde{w}$ is the vector of wage-bill fractions for the different human-capital categories. We then estimate (5) using nonlinear least squares, with the labor quality term $q_{\theta}^{\text{mult,wage}}(\tilde{w})$ multiplying the total wage bill w. The regression equation is

$$\log y = \beta_0 + \beta_l \log\left(w \cdot q_{\theta}^{\text{mult,wage}}(\tilde{w})\right) + \beta_k \log k + e.$$

[10] Value added may be formed from sales by subtracting materials costs but not the wage bill. Thus, the wage bill does not, in an accounting sense, enter the calculation of value added.

There is also a translog specification equivalent to (6).

2.3 Productivity dispersion

Although we do discuss estimates of production function parameters such as β_0, β_l, β_k and θ, our primary focus is on total factor productivity, or the residual e in (1). The parameters such as θ can be economically large and statistically significant despite the dispersion in e, the key puzzle to understand about productivity, remaining large. We focus on several related measures of the dispersion of e.

Productivity dispersion is intimately related to $R^2 = 1 - \frac{\mathrm{Var}(e)}{\mathrm{Var}(\log y)}$. One attempt to explain productivity dispersion is to add observables to the model to see how much residual productivity dispersion declines. If a new variable reduces productivity dispersion, it will also increase the statistical fit of the regression. The change in statistical fit, R^2, from adding a single new regressor z to the production function (1) estimated by ordinary least squares is

$$\Delta R^2 = \left(1 - R^2_{\text{base}}\right)\left(\text{partialcorr}\left(\log y, z \mid \log l, \log k\right)\right)^2, \tag{8}$$

where $\text{partialcorr}\left(\log y, z \mid \log l, \log k\right)$ is the partial correlation between output $\log y$ and the new input z once the non-quality adjusted inputs, $\log l$ and $\log k$, are controlled for. To compute a partial correlation, one separately regresses $\log y$ and z on $\log l$ and $\log k$ and then forms the simple correlation of the residuals from the $\log y$ and z regressions. Equation (8) indicates that a variable will add a lot of explanatory power to a regression if it is correlated with the dependent variable but is not so correlated with the other independent variables.[11]

Maximizing R^2 is the least-squares criterion. We also report our findings in terms of $\frac{\mathrm{sd}(e)}{\mathrm{sd}(\log y)} = \sqrt{1-R^2}$, the ratio of the standard deviation of productivity to the standard deviation of log value added. If input quality explains why some firms produce more outputs with the same inputs, then labor quality should decrease $\frac{\mathrm{sd}(e)}{\mathrm{sd}(\log y)}$.

Our previous two criteria work with the logged instead of the unlogged levels of productivity. Our preferred measure of productivity dispersion in unlogged levels is q_{90}/q_{10}, where q_{90} is the 90th quantile of TFP in levels $\exp(e)$ and, likewise, q_{10} is the 10th quantile of $\exp(e)$. q_{90}/q_{10} is the ratio of outputs for the 90th quantile and 10th quantile firms, if those firms had used the same inputs. For each regression we report R^2, the ratio of productivity dispersion relative to total-output dispersion $\frac{\mathrm{sd}(e)}{\mathrm{sd}(\log y)}$, as well as q_{90}/q_{10}, the non-logged level of productivity dispersion.

2.4 Productivity dispersion decline benchmarks

There is no absolute metric for whether any given decline in productivity dispersion is large or small. First, we benchmark the productivity declines from adding human-capital measures against the decline in productivity

[11] The R^2 from nonlinear least squares (NLS) is not guaranteed to be between 0 and 1 (the derivation for OLS uses the first order conditions of OLS to set the sample covariance of the residual and the predicted dependent variable to 0.) We define R^2 for NLS to be $1 - \frac{\mathrm{Var}(e)}{\mathrm{Var}(\log y)}$.

dispersion from adding firm growth and firm age. Baldwin (1995) and others show that firms that are more productive will on average have larger rates of employment growth. Cabral and Mata (2003) and others show that older firms tend to be more productive. Because of the prior literature relating productivity to growth and firm age, there is are a priori reasons to suspect that including firm age and employment growth will decrease productivity substantially.

We add firm age and growth as observed components of productivity, as in

$$\log y = \beta_0 + \beta_l \log l + \beta_k \log k + \beta_{\text{DHgrowth}} r_{\text{DHgrowth}} + \beta_{\text{firmage}} \log\left(r_{\text{firmage}}\right) + e.$$

We use the Davis and Haltiwanger (1992) measure of firm-employment growth, which ranges from -2 to 2 instead of -1 to 1 to account for firm entry and exit.[12] We see how much residual productivity dispersion declines after accounting for firm age and growth. This decline provides a benchmark for the decline in dispersion from controlling for input quality.

Another approach for benchmarking compares our productivity-dispersion decline to another decline that has shown to be important in the literature. Syverson (2004) regresses productivity dispersion (the interquartile range) in a local geographic market on a measure of the demand density (a proxy for product-market competition) in that market. He finds that a "one-standard-deviation increase in logged demand density implies a decrease in expected productivity dispersion by approximately 0.042 log points—roughly one-seventh of the mean dispersion and over one-fourth of its standard deviation." Syverson's measure of dispersion is the interquartile range of log TFP. We will compare our productivity declines from adjusting for input quality to those from Syverson from varying local-market competition.

2.5 Simultaneity bias

Marschak and Andrews (1944) introduce the endogeneity concern that more productive firms may use more inputs, leading to overestimating the input elasticities. Griliches and Mairesse (1998) argue that traditional methods of correcting regressions for endogeneity, panel data and instrumental variables, work poorly for production-function estimation because of measurement error (panel data) and data availability (instruments). Following the recent literature, we use investment to correct for input endogeneity using the Olley and Pakes (1996) estimator.[13]

The Olley and Pakes estimator decomposes productivity into what the model labels true productivity ω and measurement error η. We have a separate paper (Fox and Smeets, 2008) where we derive three empirical

[12] We use the log of firm age as firm age can have some extreme outliers (hundreds of years old) in Denmark.

[13] Ackerberg, Caves and Frazer (2007b) provide corrections to theoretical assumptions needed for the Olley and Pakes estimator to be consistent. Ackerberg et al. also introduce a new estimator that may be consistent when the labor variable is a dynamic variable, i.e. when firm tenure contributes to productivity, like in the human-capital specification in this paper. Our experiments with the Ackerberg et al. estimator show that using it does not change our conclusions about productivity dispersion.

checks that the proxy input estimators such as Olley and Pakes, Levinsohn and Petrin (2003) and Ackerberg, Caves and Frazer (2007b) should pass if the estimator is likely to be consistent.[14] The three tests are: 1) Treating ω as observable should decrease residual productivity dispersion (increase R^2) substantially; 2) η should not be correlated over time at the same firm; and 3) ω but not η should be correlated with real outcomes such as firm growth.[15] The Olley and Pakes estimator and its peers fail our tests on the Danish data and also the Chilean data that has been used in many productivity papers, including Levinsohn and Petrin and Ackerberg et al. The likely problem is that the scalar-unobservable assumption (investment is strictly increasing in true productivity ω) critically required by proxy-input estimators seems to be too strong.

We present production-function estimates using the Olley and Pakes estimator because this is the most commonly-used procedure in the literature, but we do not take a strong stand that they are likely to be consistent, under this or other data we have examined. Still, these estimators may improve the estimates of the production-function parameters, compared to using no correction. As we show empirically in Fox and Smeets, the Olley and Pakes decomposition of $e = \omega + \eta$ into true productivity ω and measurement error η is particularly empirically implausible, so we do not attempt to purge productivity of measurement error.[16]

Even our input-quality measures are imperfect. A standard matching model suggests that inputs and firms should assortatively match, if firm productivity and input quality are complements. High-ability workers should be at firms with high productivities. If so, a standard omitted-variable-bias story suggests that the parameter estimates on the human-capital variables should be biased upwards: there is a positive correlation between human capital and the true error term, productivity. Recall that equation (8) suggests that the decline in productivity dispersion from adding a variable to a regression involves the partial correlation of the new regressor with the dependent variable. If assortative matching between firms and workers increases this partial correlation, the decline in productivity dispersion from adding human-capital variables will be overstated. Therefore, this bias in the parameter estimates works against finding that the decline in productivity dispersion is small, which will be our eventual empirical finding.

3 Data overview

We start with detailed panel data on all Danish citizens for 1980–2001. These data provide us general human capital (experience, schooling), firm-specific and industry-specific human capital (firm tenure, industry tenure) as well as the monthly salary for each worker. We are careful with measuring firm tenure because of changes

[14] We experimented with the Levinsohn and Petrin (2003) procedure to proxy for productivity using capital and materials inputs. We defined materials as total sales minus value added. However, we have data on total sales and hence materials for a small sample of firms. As this sample is highly selected, we do not report the Levinsohn and Petrin estimates.

[15] Our three checks can be seen as overidentifying moments in a GMM framework. A failure of one or two of the three checks may be explained away, but it is hard to explain how the assumptions of the model of Olley and Pakes could be true if all three checks fail.

[16] Van Biesebroeck (2008) compares the Olley and Pakes (1996) estimator to four alternatives for dealing with endogeneity bias in production functions. He shows all five estimators give remarkably similar estimates of firm productivity, e in (1).

in firm identification codes. The underlying data for these variables come from government records and not subjective self-reports, like in US publicly-available microdata. Thus, we feel that our data on worker characteristics are of higher quality than any found in the United States. We aggregate our human-capital measures to the firm level to construct our labor-quality measures, as in (3) and (7). We also compute the total number of workers as well as several wage-bill measures.

We then merge the firm-level human-capital measures with data on value added, physical capital and investment.[17] These data come from a credit-rating agency, for the year 2001. More details on the data are found in the appendix.

Denmark is a small open economy, so there are not many distinct firms in narrowly-defined industries. We strive to balance the competing needs to have more observations for precise statistical inference and to allow heterogeneity in the production functions for firms in different industries. We consider a medium level of aggregation because we include many detailed measures of human-capital variables and therefore need a lot of observations per regression. We perform separate regressions for eight industries: furniture, food and beverages, publishing and printing, fabricated metals, machinery and equipment, hotels and restaurants, construction of complete structures, and advertising. To alleviate some forms of heterogeneity, we include fixed effects at the five-digit industry level in each regression.

Table 1 lists summary statistics for four of our eight industries: those with production-function estimates in Tables 2–5. Value added and inputs vary a lot across firms. Importantly, the human-capital measures vary a lot across firms. As there is variation across firms in the sample, equation (8) suggests it is a priori possible that adding human-capital quality measures to a production function will increase the R^2 and hence reduce the dispersion in measured productivity.

4 Production function estimates

The paper's focus is on productivity dispersion, which arises from the dispersion of the residuals from these regressions. Before discussing productivity dispersion in detail, we will describe the production-function estimates in order to provide context. As a warning, we do not feel that the production-function parameter estimates are robust empirical findings. Our results on productivity dispersion are very robust across functional-form choices. All of our standard errors allow for heteroskedasticity.

Tables 2–5 report production-function estimates for four industries. Table 2 covers the food and beverages industry. Column 1 is a base specification, with just the number of workers for the labor input. The coefficient on labor is 0.81 and the coefficient on physical capital is 0.21, resulting in an estimate of a small increasing returns to scale.[18] The R^2 from the base regression is 0.862.

[17] We do not observe measures of inputs other than labor and physical capital and we do not observe sales for many firms.

[18] As the dependent variable is sales and not physical output, Klette and Griliches (1996) suggest that the returns to scale will be biased

In Column 2, we begin to account for labor quality. Column 2 uses the numbers of college and non-college workers as separate inputs. The coefficient on the number of skilled workers is 0.34, and the coefficient on the unskilled workers is 0.52. The coefficient on physical capital declines to 0.17. R^2 increases by only 0.01, from 0.862 to 0.872.

Column 3 shows the estimates from (5) with the multiplicative / Griliches (1957) labor quality term, (3). The coefficient on female is -0.537, which can be interpreted as saying that a firm with 10% more of its workforce being women will have $1 - 0.537 \cdot 0.10 = 0.9463$ or 95% of the total labor inputs $l \cdot q_\theta(x)$ as another firm with the same number of workers, l. Schooling is one of our main measures of general human capital. The coefficient of 3.1 on the fraction of college-educated workers says that a firm with 10% higher fraction of college-educated workers (as opposed to the excluded category, workers who completed high school or below) will have 31% more labor inputs. The coefficient is statistically significant, but the coefficients on the fraction of workers with community college and vocational degrees are not statistically distinct from 0. The coefficient on community college is large in terms of its economic magnitude, however.

One of our data advantages is that we can construct detailed labor-history measures using our worker panel data. Total experience in the labor market is exactly computed at the worker level from government records (since 1964). With no concern about topcoding experience, we enter experience as the mean level of experience of workers at the firm, mostly to save space in the tables. A firm whose workforce has an extra 10 years of labor-market experience will have 13% more labor inputs.

We next look at firm tenure and industry tenure in column 3. These approximate firm- and occupational-specific human capital. The measures are the percentage of workers in each tenure bin, and all coefficients should be evaluated relative to the residual category, newcomers with 0 years of tenure. We find a firm with 10% more workers with 1–2 years of tenure instead of newcomers has $0.10 \cdot 0.742 = 7.4\%$ more labor inputs, a potentially large effect. Because of the large standard errors, the coefficients for the firm-tenure categories are mostly consistent with a large, one-time training cost for newcomers.[19] Three of the four industry-tenure coefficients are negative. The largest negative coefficient is statistically distinct from 0. The R^2 from the multiplicative-labor-quality specification is 0.886. Overall, we have several statistically-significant coefficients and many coefficients with economically-large magnitudes. Our finding of a small decrease in productivity dispersion from labor-quality controls will not be due to economically-small or statistically-insignificant estimates of human-capital production-function parameters.

Column 4 of Table 2 uses the Welch (1969) additive-labor-quality function. While the coefficients are not directly comparable in magnitude to those using the multiplicative specification in column 4, several of the

downwards. This bias could be offset by other biases such as the usual bias that more productive firms use more inputs, which tends to bias the returns of scale parameters upwards.

[19] A potential "training cost" pattern of coefficients may also reflect a measurement issue: workers hired during the year at a growing firm who are mistakenly counted as working the entire year. This is an issue for growing firms and not firms with simply higher levels of turnover. We reran the labor-quality specification by adding the past 5-year firm-employment growth using the Davis and Haltiwanger (1992) measure and the extra regressor increases the magnitude of the firm-tenure coefficients, which goes against the growing-firms explanation.

coefficients do change sign. For example, the industry-tenure coefficients are all positive, while experience has a negative coefficient. Because the estimated signs of the labor-quality coefficients are sensitive to the functional form for the labor-quality function, we do not view the signs of the point estimates of the labor-quality terms as robust findings.

The robust finding across columns 3 and 4 is the R^2: 0.886 in column 3 and 0.885 column 4. We return to productivity dispersion in the next section.

Column 5 attempts to adjust for the quality of the workforce by using the wage bill. The coefficient on the total wage bill is higher than the corresponding input elasticity when the number of workers is used instead, in column 1. The coefficient on physical capital decreases. The R^2 increases to 0.887, which is slightly higher than the R^2 for the specifications with the detailed human-capital measures.

Column 6 uses the specification that combines human-capital and wage-bill data, (7). Labor quality uses the multiplicative form. The R^2 increases to 0.894, which is to be expected for the specification using the most data. Most of the coefficients are not statistically distinct from 0 at the 95% level.[20]

Interpretations of the parameters require a convincing argument that the labor inputs are uncorrelated with the error term, productivity.[21] Studies that do not correct for endogeneity argue that more productive firms employ higher-quality workers (Haltiwanger et al., 2007).[22] As discussed above, we adopt the advice of Griliches and Mairesse (1998) and use the Olley and Pakes (1996) estimator, which corrects for the correlation of the inputs with productivity by proxying for productivity with investment. A panel-data moment condition is also used for identification.[23] We use the multiplicative functional form for labor quality, (3). In column 7 of Table 2, the Olley and Pakes estimator raises the coefficient on physical capital (from 0.15 to 0.25), which is often thought to be too low when estimated via OLS. Comparing with column 3, correcting for endogeneity also changes some of the human-capital point estimates, but not many of the signs of the point estimates. The R^2 listed in column 7 reports $1-\frac{\text{Var}(\omega+\eta)}{\text{Var}(\log y)}$, where ω is the true productivity and η is the measurement error according to the model in Olley and Pakes. The R^2 decreases to 0.837.[24] The R^2 decreases with the Olley

[20] For the wage-bill specification, our statistics package Stata occasionally does not report standard errors for two parameters. Reviewing the method for computing standard errors in Stata's user manual, we suspect that this is because the matrix of gradients of the production function (not the objective function), stacked across the statistical observations, is singular at the converged point estimates. We have verified that the point estimates themselves are true local minima by experimenting with several sets of starting values for the nonlinear-optimization routine. We also experimented with using the bootstrap to construct standard errors. Those standard errors are reported in footnotes to the table.

[21] Measured productivity is a residual and will be uncorrelated with included inputs in a linear regression, by the OLS first-order conditions. However, true productivity may be correlated with inputs.

[22] We do not take a strong stand that our estimates of the human-capital parameters are causal production-function estimates. Our main focus is on the dispersion of the productivity residual e, which seems to be relatively invariant to the method used to estimate the production function.

[23] We use a subset of the data because we need firms with nonmissing investment data in both 2001 and the previous year, 2000. We use a fifth-order polynomial in firm age, physical capital and investment in the Olley and Pakes first stage. In the second stage, we use nonlinear least squares to estimate equation (34b) in Ackerberg, Benkard, Berry and Pakes (2007a). We also include firm age as a firm state variable, like in Olley and Pakes.

[24] The sample has decreased because of missing investment data. We estimated the multiplicative labor-quality production function using the sample with non-missing investment data but without the Olley and Pakes adjustment for simultaneity bias. The R^2 is nearly the same as with the full sample in column 3.

and Pakes specification, in large part because some explanatory power from the measured inputs like k and l is now shifted to the Olley and Pakes model's true-productivity term (ω) when that term is proxied by a combination of physical capital, firm age and investment. Under the interpretation from Olley and Pakes's model, the dispersion in measurement error η is almost the same as the dispersion in e without endogeneity correction. The measurement-error interpretation for η contradicts much of the empirical literature, which shows that e is predictive of real outcomes like firm growth and firm exit.[25]

Column 8 is a benchmark regression. Some of the productivity literature finds that firms that are older and that firms that have recently grown quickly are more productive. We have a direct measure of firm age and can construct the past five years of firm employment growth, using the Davis and Haltiwanger (1992) measure. We include these measures as extra regressors in a standard Cobb-Douglas regression, (1). In column 8, we find that these measures are not very predictive of firm output in the food and beverages industry. The R^2 barely increases over the base case in column 1.

Table 3 reports the same set of eight production-function estimates for the publishing-and-printing industry. The returns to scale are higher than in food and beverages: column 1 has a return to scale of almost 5%. The R^2 with no labor-quality measures is 0.836, which increases to only 0.837 by adding college and non-college workers as separate inputs. The R^2 with either the additive or multiplicative labor-quality measures is 0.862. Unlike food and beverages, the signs of the human-capital coefficients are mostly the same for the multiplicative and additive functional forms. The R^2 of the pure wage-bill specification is 0.868, higher than the R^2 from the human-capital specifications. The returns to scale decrease in the wage-bill specification. The specification in column 6 that combines wage-bill and human-capital data has an R^2 of 0.875. As in food and beverages, the R^2 for the Olley and Pakes estimator decreases, as some explanatory power from the measured inputs is transferred to the ω term.

Table 4 reports another set of estimates, this time for the fabricated-metals industry. The R^2 for the base case is 0.726. Interestingly, most specifications find a decreasing return to scale. There is a strong level of agreement between the multiplicative and additive specifications about the signs of the point estimates for the human-capital coefficients. A robust empirical finding is that both the additive and multiplicative labor-quality adjustments give an R^2 of 0.740. The specification with the wage bill has an R^2 of 0.750 and the specification with both the wage-bill and human-capital data gives an R^2 of 0.758. A change in R^2 of only $0.758 - 0.726 = 0.032$ from adding detailed human-capital and wage-bill variables is small.

Table 5 presents estimates from the furniture industry. Like in food and beverages, many of the coefficients on the human-capital variables in column 3 are statistically significant. Despite some of the human capital variables having statistically-significant and economically-large point estimates, the same patterns about R^2 as

[25]Column 7 also lists the standard deviations of ω and η. Most of the productivity dispersion is attributed to η, not ω. In Fox and Smeets (2008), we show that, for example, η is autocorrelated and η predicts real outcomes like firm exit and growth, which it should not do under a measurement-error story. Therefore and in agreement with most of the empirical literature, we do not feel that the main explanation for productivity dispersion is that it reflects high measurement error from the η term.

in other industries arise. The base case R^2 is 0.812, it increases to 0.843 with detailed human-capital measures, 0.847 with the wage bill, and 0.858 with both the wage-bill and detailed human-capital measures.

For conciseness, we do not report the parameter estimates for the translog production functions or for the other four industries.

5 Productivity dispersion and input quality

Table 6 is the main result of the paper. The table reports the R^2, the standard deviation of log TFP (e), the ratio $\frac{\mathrm{sd}(e)}{\mathrm{sd}(\log y)}$, and our measure of productivity dispersion in levels of output instead of logs, q_{90}/q_{10}. For each of our eight industries, the first row is a baseline specification with the usual measure of labor, the number of workers. The second row uses simple labor-quality measures previously found in the literature: the (log) number of workers with college degrees and the (log) number of workers without college degrees, as separate inputs. The third row is perhaps our main specification: the estimation of (5) using the detailed general- and specific-human-capital measures. The fourth row tests the robustness of the findings on productivity to the choice of functional forms for labor quality.

The fifth row replaces the number of workers with the wage bill. Wages may proxy for worker labor quality in a competitive labor market. On the other hand, using wages may introduce reverse-causality problems because of profit sharing or effort-raising work practices: more productive firms may pay higher wages to equivalent workers. The sixth row uses both the wage-bill and human-capital measures. Finally, the seventh row is a benchmark, where we use data on firm age and recent firm-employment growth. All specifications include fixed effects for five-digit sub-industries. We do not include the Olley and Pakes estimates because the sample is smaller because of missing investment data, and so the productivity dispersion is not directly comparable.

Consider food and beverages. Including only the number of workers gives a R^2 from (1) of 0.86. Therefore, the standard deviation of the residual e is $\sqrt{(1-0.86)\cdot 1.52^2} = 0.57$, where 1.52 is the standard deviation of log value added. The ratio $\frac{\mathrm{sd}(e)}{\mathrm{sd}(\log y)}$ is 0.37: productivity dispersion is 37% of the dispersion in log value added. Also, the ratio of the 90th quantile of $\exp(e)$ to the 10th quantile of unlogged TFP is 3.48. A firm at the 90th quantile produces 3.48 times the output as a firm at the 10th quantile, for the same inputs. Although all the dispersion measures are listed in Table 6, for conciseness our discussion in the text will focus on q_{90}/q_{10} as our main measure of productivity dispersion. q_{90}/q_{10} is more related to economic outcomes as it involves productivity in unlogged levels and does not normalize the measure by the dispersion of value added.

Continuing with food and beverages, we now explore the reductions in productivity dispersion from including input-quality measures. Disaggregating workers into separate college and noncollege inputs, as is sometimes done in the literature, decreases q_{90}/q_{10} from 3.48 to 3.31. Our most important specifications are the ones that use the detailed human-capital measures. q_{90}/q_{10} is 2.98 for the multiplicative labor-quality functional form and 3.12 for the additive functional form. An alternative to using human-capital measures is to use the wage bill. With the wage bill, q_{90}/q_{10} is 3.00. The wage bill gives lower productivity dispersion than the human-capital measures. q_{90}/q_{10} is also 3.00 with both the wage-bill and human-capital measures. Finally, the benchmark of firm growth and firm age contributes very little to decreasing productivity dispersion: q_{90}/q_{10} is 3.38 when these variables are added to the base case.

We also estimated all seven specifications using a translog production function. Table 6 also lists these results. Remarkably, the estimates of productivity dispersion as measured by $\mathrm{sd}(\log e)$ are quite similar whether the production function is a Cobb-Douglas or a translog. The measure q_{90}/q_{10} is more sensitive to the functional form for the production function. Adding additional nonlinear terms to a Cobb-Douglas can only weakly increase R^2 and hence will often (although it is not a theorem) decrease productivity-dispersion measures such as q_{90}/q_{10}. For food and beverages, Table 6 shows that these extra polynomial terms in most cases decrease q_{90}/q_{10} some, but not necessarily by a large amount. The decline in q_{90}/q_{10} from adding employment growth and firm age is slightly larger than in the Cobb-Douglas case.

There are seven other industries listed in Table 6. The pattern of productivity dispersion is qualitatively the same in all the eight industries we looked at. Adding college and noncollege workers as separate inputs decreases dispersion hardly at all. Adding detailed human-capital controls decreases productivity dispersion by more. In all industries, the productivity dispersion is roughly invariant to whether a multiplicative or additive labor-quality functional form is used. The wage bill is potentially a more accurate measure of input quality than the detailed human-capital measures. Indeed, the wage-bill specification usually gives less dispersion than the human-capital specifications. Unsurprisingly, the specification with both wage-bill and human-capital data decreases dispersion the most.

Table 6 also provides estimates of productivity dispersion for the translog production functions. Almost all of the $\frac{\mathrm{sd}(e)}{\mathrm{sd}(\log y)}$ ratios of productivity dispersion are numerically identical to those from the Cobb-Douglas production functions; any reported differences are mostly due to rounding. The estimates of q_{90}/q_{10} vary slightly more. In unreported results, we have shown that the main results about productivity hold when using a CES production function. Altogether, our results about productivity dispersion are mostly invariant to the functional form of the production function as well as the functional form of the labor-quality function. This finding about productivity dispersion contrasts with the signs of the parameter estimates of the production functions, which we argued above are sometimes but not always robustly estimated across functional forms.

How large are these productivity-dispersion declines? Most of the previous empirical literature has studied manufacturing, because of data availability and because production processes are likely more homogeneous across firms in manufacturing. Five of our industries are comfortably in manufacturing, and construction has some similarities with manufacturing. Averaging across these six manufacturing industries, productivity dispersion as measured by q_{90}/q_{10} declines from a mean of 3.74 (with only the number of workers) to a mean of 3.36 with the specification with both human-capital and wage-bill data. Even with the most detailed input quality measures, a firm at the 90th quantile of productivity produces 3.36 times the output of a firm at the 10th quantile. This is a tremendous amount of heterogeneity for firms in the same industry. Given this large amount of remaining productivity dispersion, our main conclusion is that not accounting for labor-input quality is not the main explanation for the tremendous amount of productivity dispersion that has been seen as one of the great puzzles by empirical economists.

Advertising as well as hotels and restaurants are both in the service sector. For hotels and restaurants, the decline in q_{90}/q_{10} is from 3.86 for the base case to 3.13 for the model with detailed human-capital and wage-bill data. For advertising, q_{90}/q_{10} starts at 3.49 for the base case and drops to 2.20 for the case with both wage-bill and human-capital data. The likely high level of human-capital intensity in advertising is a likely explanation for the drop. Another possibility for the larger decline in advertising is that the production technology is more heterogeneous in the service sector, and the human-capital and wage-bill measures pick up heterogeneity across firms in the production function. This is somewhat consistent with findings from other data that service industries have higher overall dispersion (Oulton, 1998).

Based on prior research, we used each firm's employment growth over the last five years and the log of firm age as a benchmark for productivity dispersion decline. Firm age and firm growth did not substantially decrease productivity dispersion in any of the eight industries. Syverson (2004) studied local demand density (a proxy for competition) and productivity dispersion and found that a one-standard deviation increase in demand density lowered the interquartile range of e by -0.042 log points. For food and beverages, the interquartile range of productivity e is 0.586 for the base case without labor-quality adjustment and 0.483 for the specification with both wage-bill and human-capital data. Syverson (2004) studied narrow geographic markets for a homogeneous product, concrete. It is not surprising that the mean level of productivity dispersion of 0.275 in his paper is half of our base value of 0.586. Starting from a higher base dispersion, the productivity-dispersion decrease from adding human-capital and wage-bill data is 0.103 log points, or equivalent to a 2.5 standard-deviation increase in demand density across local markets in the concrete industry. Our interpretation is that adding human-capital variables produces productivity-dispersion declines roughly on the same order of magnitude as within-sample changes in demand densities.

6 Conclusions

Since at least Griliches (1957), economists have speculated that productivity dispersion may arise because firms use inputs of varying qualities. We study labor inputs in part because physical capital is already quality adjusted, as physical capital is usually measured in monetary units. By contrast, researchers often use the number of workers for the labor input. We use detailed data on all Danish citizens to construct human-capital measures at the firm level. Human-capital inputs do vary across companies in Denmark and our production-function parameter estimates show human-capital inputs raise firm output considerably. For some industries, the human-capital coefficients are statistically precisely estimated.

Adding these quality-adjusted inputs does not dramatically decrease within-industry productivity dispersion.

For all of our six manufacturing industries, the decline in productivity dispersion was not large. Averaging across the manufacturing industries, productivity dispersion as measured by q_{90}/q_{10} declines from a mean of 3.74 (with only the number of workers) to a mean of 3.36 with the specification with both human-capital and wage-bill data. Using our most detailed input-quality measures, the 90th-quantile-TFP firm produces 3.36 times the output of the 10th quantile firm, for the same inputs. The high remaining productivity dispersion, with the best available input-quality data researchers are likely to have access to, suggests that input-quality dispersion is unlikely to be the main factor explaining why firms in the same industry have different levels of output. On the other hand, the decline is somewhat larger for the human-capital-intensive, service-sector-industry advertising. Still, the remaining productivity dispersion is large even in advertising.

The decline in productivity dispersion from adding controls for firm age and firm growth, two measures emphasized in the literature, was barely detectable. Compared to this benchmark, the decline in productivity dispersion from input quality is large. The decline in productivity dispersion is roughly the same order of magnitude as the competitive effects studied in local geographic markets by Syverson (2004). Perhaps input quality is one of a string of items that together combine to explain productivity dispersion. Still, our main conclusion is that economists should cease listing (labor) input quality as a major item in the list of possible explanations for productivity dispersion. The idea has a lot of theoretical appeal, but simply does not seem to be true.

Returning to an issue we raised in the introduction, our results suggest that productivity represents some attribute of a firm that cannot easily be bought and sold on the market for inputs. Possibilities include management quality, business strategy, the appropriate use of new technologies and heterogeneous production technologies. If productivity cannot be traded, then the performance of product markets may be as important for economic efficiency and aggregate productivity growth as the performance of input markets.

A Danish labor and accounting data

We use accounting data for capital, value added and investment. The accounting data come from Købmanstandens Oplysningsbureau (KØB), a Danish credit-rating agency. The accounting data are an unbalanced panel that roughly covers the period 1995–2003 and uses each firm's proprietary accounting period. We rescale the accounting variables to a twelve-month, calendar-year basis. We look at the year 2001 to maximize the number of firms with complete calendar-year data.

We use value added as a measure of output and our measure of physical capital is tangible assets net of depreciation. Value added is reported for many more firms than total sales, perhaps because of the role of value added in value-added taxes. We disregard firms that lack rescaled accounting information on valued added and fixed assets for a twelve-month period. For the labor input, we count the total number of workers in IDA, which is described below. Firm age is directly reported in the accounting data. We include the log of firm age in some specifications.[26]

To construct labor-quality variables, we use the Danish Integrated Database for Labor Market Research (IDA), one of the central registers of Statistics Denmark. IDA combines several types of data. One dataset provides information at the individual level on demographics (age, sex, marital status, family status) and schooling for all Danish citizens for 1980–2001. Each individual is given a unique identification number that can be further used for matching with the other datasets of IDA. Another IDA dataset's unit of observation is an individual's job. It contains information on individual labor earnings, some other variables and the number of years of labor market experience. Labor market experience is computed since 1964 by Statistics Denmark.

Both full- and part-time jobs are included, but in the rare case of a worker with three or more jobs, only the primary and secondary jobs are reported. The data also contain a unique identification number for each job's establishment. IDA's establishment dataset provides a firm identification number that can be use for matching with other firm-level data.

We use IDA for 1980–2001 to compute labor-market-history variables such as firm tenure and industry tenure. We compute firm tenure as the number of years a worker has been attached to a given firm. As we are concerned with spurious changes in firm identification codes, a worker's tenure is reset to zero only if both his firm and establishment identifiers change at the same time. We construct industry tenure using the following eight broad sectors: (1) agriculture and mining, (2) manufacturing, (3) construction and transport, (4) retail, hotels and restaurants, (5) finance, real estate and R&D activities, (6) public sector, (7) private households and extraterritorial activities and (8) others. These sectors encompass all Danish firms and are not equivalent to the industries for our estimation sample.

Industry is recorded at the establishment level. For our regressions, a multi-establishment firm's industry is the weighted (by number of workers) modal establishment industry.

All inputs are constructed at the firm level. We construct firm-level fractions of workers who have a given characteristic, say a college degree or 6–9 years of firm tenure. The intervals are simple to interpret as each measure is a fraction between 0 and 1. The intervals allow us to examine nonlinearities, and they handle topcoding from not observing firm and industry tenure for spells starting before 1980.

We estimated production functions for two samples: all firms with nonmissing variables and a sample with outliers removed. We are worried about possibly non-classical measurement error in the accounting data, so we removed the firms in the top and bottom 1% of the ratios of output to labor and also physical capital to labor. Removing these outliers increases the base R^2's substantially, but does not change the ΔR^2's from adding labor quality much. We report specifications with the outliers removed, but our main conclusions about ΔR^2's are similar if we include the outliers.

[26] We construct investment from the accounting data in order to control for the endogeneity of the labor input using the Olley and Pakes (1996) approach. Investment is computed using the formula $i = k_{2001} - (1-\delta)\,k_{2000}$, where δ is the depreciation rate. Investment cannot be missing and firms must be present in both 2000 and 2001. The accounting data report δk_{2000}, which we use to back out δ.

References

Ackerberg, Daniel, C. Lanier Benkard, Steven Berry, and Ariel Pakes, "Econometric Tools for Analyzing Market Outcomes.," in James J. Heckman and Edward E. Leamer, eds., *Handbook of Econometrics*, Vol. 6A, Elsevier, 2007.

_ , Kevin Caves, and Garth Frazer, "Structural Estimation of Production Functions," 2007. working paper.

Baily, Martin Neil, Charles Hulten, and David Campbell, "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity: Microeconomics*, 1992.

Baldwin, John R., *The Dynamics of Industrial Competition: A North American Perspective*, Cambridge University Press, 1995.

Bartelsman, Eric J. and Mark Doms, "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, September 2000, *38* (3), 569–594.

Bernard, A.B. and J.B. Jensen, "Exporters, Jobs, and Wages in US, Manufacturing: 1976-1987," *Brookings Papers on Economic Activity: Macroeconomics*, 1995.

Bloom, Nick and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, November 2007.

Cabral, Luis M. B. and José Mata, "On the Evolution of the Firm Size Distribution: Facts and Theory," *American Economic Review*, November 2003, *93* (4), 1075–1090.

Christensen, Laurits R., Dale W. Jorgenson, and Lawrence J. Lau, "Transcendental Logarithmic Production Functions," *The Review of Economics and Statistics*, February 1973, *55* (1), 28–45.

Davis, Steven J. and John C. Haltiwanger, "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics*, August 1992, *107* (3), 819–863.

Denison, Edward Fulton, *The Sources of Economic Growth in the United States and the Alternatives Before Us*, Committee for Economic Development, 1962.

Dhrymes, Phoebus, *The Structure of Production Technology: Productivity and Aggregation Effects*, Edward Elgar, 1995.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, March 2008.

Fox, Jeremy T. and Valérie Smeets, "Three Empirical Checks for the Validity of Olley and Pakes Productivity Estimators," 2008. working paper.

Frazer, Garth, "Heterogeneous Labour in Firms and Returns to Education in Ghana," 2006. working paper.

Griliches, Zvi, "Specification Bias in Estimates of Production Functions," *Journal of Farm Economics*, February 1957, *39* (1), 8–20.

_ **and Jacques Mairesse**, "Production Functions: The Search for Identification," in S. Strøm, ed., *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, Cambridge University Press, 1998.

Haltiwanger, John, Julia Lane, and James Spletzer, "Wages, Productivity, and the Dynamic Interaction of Businesses and Workers," *Labour Economics*, 2007.

Hellerstein, Judith and David Neumark, "Production Function and Wage Equation Estimation with Heterogeneous Labor: Evidence from a New Matched Employer-Employee Data Set," in "Hard to Measure Goods and Services: Essays in Honor of Zvi Griliches," University of Chicago, 2006.

Jorgenson, D.W., F.M. Gollop, and B. Fraumeni, *Productivity and US Economic Growth*, Harvard University Press, 1987.

Katayama, Hajime, Shihua Lu, and James R. Tybout, "Firm-Level Productivity Studies: Illusions and a Solution," November 2006. working paper.

Klette, Tor Jakob and Zvi Griliches, "The Inconsistency of Common Scale Estimators when Output Prices are Unobserved and Endogenous," *Journal of Applied Econometrics*, 1996, *11*, 343–361.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, *70*, 317–341.

Maksimovic, V. and G. Phillips, "The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains?," *The Journal of Finance*, 2001, *56* (6), 2019–2065.

Marschak, Jacob and William H. Andrews, "Random Simultaneous Equations and the Theory of Production," *Econometrica*, July-October 1944, *12* (3/4), 143–205.

Mulligan, C.B. and X. Sala-I-Martin, "Measuring Aggregate Human Capital," *Journal of Economic Growth*, 2000, *5* (3), 215–252.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, November 1996, *64* (6), 1263–1297.

Oulton, N., "Competition and the dispersion of labour productivity amongst UK companies," *Oxford Economic Papers*, 1998, *50* (1), 23–38.

Schoar, A., “Effects of Corporate Diversification on Productivity,” *The Journal of Finance*, 2002, *57* (6), 2379–2403.

Syverson, Chad, “Market Structure and Productivity: A Concrete Example,” *Journal of Political Economy*, 2004, *112*, 1181–1222.

Van Biesebroeck, Johannes, “Wages Equal Productivity: Fact or Fiction?,” June 2007. working paper.

_ , “The Sensitivity of Productivity Estimates: Revisiting Three Important Debates,” *Journal of Business and Economic Statistics*, July 2008, *26* (3), 321–338.

Welch, Finis, “Linear Synthesis of Skill Distribution,” *The Journal of Human Resources*, Summer 1969, *4* (3), 311–327.

TABLE 1 - Summary Statistics by 2-digit Industries†

Variables	Food and Beverages		Publishing and Printing		Fabricated Metals		Furniture	
	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.
Value added	101,400	375,900	37,300	95,900	17,300	28,800	26,300	82,900
Log value added	10.1	1.5	9.5	1.2	9.2	1.0	9.4	1.1
Capital	146,900	594,300	34,700	97,800	15,400	39,100	27,100	100,800
Labor	223.6	864.0	77.2	151.9	51.1	79.9	79.1	231.2
Firm age	28.0	32.3	27.7	31.5	24.0	21.7	29.3	27.2
DH Growth 1996-2001	0.32	0.67	0.27	0.60	0.27	0.57	0.16	0.53
Experience	14.4	4.6	15.3	4.3	15.9	3.3	15.7	3.4
Female (%)	14.4	4.6	41.2	16.2	15.5	12.5	30.7	17.7
College & master (%)	6.2	6.2	14.7	14.8	5.9	5.6	5.9	6.2
Community college (%)	5.5	5.5	3.4	4.3	5.6	4.8	4.6	4.8
Vocational (%)	44.8	12.7	58.8	16.6	56.5	13.8	50.8	13.6
Firm tenure 1 to 2 years (%)	24.5	13.0	26.2	14.3	25.0	14.1	22.7	11.1
Firm tenure 3 to 5 years (%)	17.9	9.8	17.6	11.7	18.5	10.9	20.2	10.8
Firm tenure 6 to 9 years (%)	11.8	9.7	11.6	10.3	13.9	10.7	13.9	10.3
Firm tenure 10 years and up (%)	17.5	14.8	20.9	16.7	20.2	15.1	21.5	15.5
Industry tenure 1 to 2 years (%)	23.2	13.9	19.6	11.9	21.1	14.0	21.5	14.6
Industry tenure 3 to 5 years (%)	17.1	10.0	16.7	10.6	18.4	10.6	19.8	12.5
Industry tenure 6 to 9 years (%)	12.4	10.0	13.3	10.0	16.7	11.5	16.3	13.4
Industry tenure 1o years and up (%)	21.3	15.6	32.3	19.3	25.7	16.6	23.8	17.8
# observations	256		277		548		254	

† Food and beverages, publishing and printing, fabricated metals and furniture are a subset of manufacturing

TABLE 2: Labor Quality Augmented Cobb Douglas Production Function
for the Food and Beverage Industry

Dep. variable: Log Value Added	(1) Number of workers		(2) College/non-college		(3) Detailed human capital measures - multiplicative form		(4) Detailed human capital measures - additive form		(5) Wage bill		(6) Detailed human capital measures - wage bill		(7) Detailed human capital measures with OP - multiplicative		(8) Benchmark - firm age and firm growth	
	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.
Log number of workers (# or wages)	0.813***	0.079			0.851***	0.075	0.865***	0.073	0.899***	0.068	0.862***	0.069	0.736***	0.173	0.797***	0.080
Log college workers			0.340***	0.062												
Log non-college workers			0.516***	0.087												
Log physical capital	0.209***	0.057	0.166***	0.051	0.149***	0.054	0.155***	0.052	0.108**	0.053	0.120**	0.052	0.250***	0.027	0.212***	0.057
Firm age (log)													-0.122**	0.050	0.030	0.038
DH growth 5 years															-0.114	0.076
Female					-0.537***	0.177	-1.061**	0.535			-0.323	0.291	-0.636*	0.347		
College & master					3.066**	1.447	6.333	4.791			1.625*	0.976	3.405	2.786		
Community college					1.347	1.357	5.116	4.112			1.329	1.191	2.476	3.637		
Vocational					0.345	0.559	0.718	1.053			0.241	0.519	0.620	1.453		
Experience					0.013	0.026	-0.011	0.041			-0.012	0.010	-0.002	0.033		
Firm tenure 1 to 2 years					0.742	1.014	-1.006	0.798			0.270	0.572	0.362	1.766		
Firm tenure 3 to 5 years					2.610*	1.502	0.896	1.644			3.498	2.259	-0.870	1.144		
Firm tenure 6 to 9 years					0.428	0.875	-1.285	1.879			0.228	[illegible]	-0.288	1.489		
Firm tenure 10 years and up					1.578*	0.980	1.000	1.604			0.698	6.800	1.542	2.606		

Industry tenure 1 to 2 years					-0.055	0.753	1.636	1.312			0.270	⌖	1.601	3.060		
Industry tenure 3 to 5 years					-1.160***	0.156	0.019	1.675			-0.896**	0.396	1.184	3.664		
Industry tenure 6 to 9 years					0.604	0.802	3.556	1.893			0.228	0.601	2.500	4.171		
Industry tenure 10 years and up					-0.339	0.427	1.536	1.808			0.332	5.940	1.841	2.824		
SD (ω)													0.584			
SD (η)													1.445			
Industry dummies	5-digit		5-digit		5-digit		5-digit		5-digit		5-digit		5-digit		5-digit	
R-squared	**0.862**		**0.872**		**0.886**		**0.885**		**0.887**		**0.894**		**0.837**		**0.865**	
# observations	256		256		256		256		256		256		177		256	

All estimations include a constant term and robust standard errors are reported. ***/**/* reports significance at 1/5/10%.

Sub-industry indicators (at the five-digit level) are included in all regressions (outside the labor quality function).

(1) Cobb Douglas

(2) Cobb Douglas with the number of workers broken down into college and non-college workers

(3) Nonlinear estimation of a multiplicative labor quality augmented Cobb Douglas production function.

(4) Cobb Douglas substituting the number of workers by the firm's monthly wage bill

(5) Nonlinear estimation of an additive labor quality augmented Cobb Douglas production function.

(6) Nonlinear estimation of a multiplicative labor quality augmented Cobb Douglas production function, using the wage bill instead of the number of workers.

(7) Nonlinear estimation of a labor quality augmented Cobb Douglas production function using the Olley and Pakes (1996) estimator to correct for input correlation with true productivity

(8) Cobb Douglas with firm age in logs, firm growth over the last 5 years using the Davis and Haltiwanger (1992) measure. Firm employment growth is calculated using the Davis and Haltiwanger (1992) measure: (xt-xt-5)/((xt+xt-5)/2)

⌖ Stata does not calculate standard errors because of likely collinearity in derivatives at the convergence point. We ran bootstrapped standard errors and they were respectively 0.297 for firm tenure 6 to 9 years and 0.499 for industry tenure 1 to 2 years.

TABLE 3: Labor Quality Augmented Cobb Douglas Production Function
for the Publishing and Printing Industry

Dep. variable: Log Value Added	(1) Number of workers		(2) College/non-college		(3) Detailed human capital measures - multiplicative form		(4) Detailed human capital measures - additive form		(5) Wage bill		(6) Detailed human capital measures - wage bill		(7) Detailed human capital measures with OP - multiplicative		(8) Benchmark - firm age and firm growth	
	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.
Log number of workers (# or wages)	0.921***	0.059			0.955***	0.060	0.963***	0.059	0.966***	0.050	0.932***	0.051	0.940***	0.111	0.922***	0.059
Log college workers			0.319***	0.052												
Log non-college workers			0.585***	0.069												
Log physical capital	0.124***	0.034	0.144***	0.035	0.0925***	0.034	0.086***	0.033	0.058*	0.031	0.074**	0.032	0.066***	0.022	0.126***	0.034
Firm age (log)													-0.091**	0.036	-0.020	0.030
DH growth 5 years															0.023	0.063
Female					-0.148	0.222	-0.835	0.786			-0.007	0.248	-0.139	0.463		
College & master					1.734**	0.765	4.145	3.302			0.434	0.471	2.515*	1.445		
Community college					1.226	1.226	3.198	3.807			0.790	0.889	0.852	1.652		
Vocational					0.427	0.466	1.190	1.234			-0.275	0.219	0.469	0.873		
Experience					0.127	0.090	0.169	0.127			0.008	0.012	0.490	1.306		
Firm tenure 1 to 2 years					-0.303	0.287	-0.979	1.254			-0.705***	0.185	1.813	1.432		
Firm tenure 3 to 5 years					-0.162	0.387	-1.001	1.512			-0.148	0.278	-0.719	0.458		
Firm tenure 6 to 9 years					-0.180	0.328	-0.261	1.353			-0.193	0.276	0.722	0.935		
Firm tenure 10 years and up					-0.410	0.257	-1.606	1.509			-0.725***	0.254	0.072	0.935		
Industry tenure 1 to 2 years					0.030	0.424	0.239	1.518			0.653	0.649	-0.818***	0.307		
Industry tenure 3 to 5 years					-0.015	0.492	0.416	1.695			-0.148	[illegible]	0.235	1.136		

Industry tenure 6 to 9 years					-0.199	0.347	-0.748	1.420			-0.193	⌖	0.251	0.957		
Industry tenure 10 years and up					-0.109	0.344	-0.563	1.295			0.665	0.693	-0.218	0.711		
SD (ω)													0.464			
SD (η)													1.199			
Industry dummies	5-digit		5-digit		5-digit		5-digit		5-digit		5-digit		5-digit		5-digit	
R-squared	**0.836**		**0.837**		**0.862**		**0.862**		**0.868**		**0.875**		**0.850**		**0.837**	
# observations	277		277		277		277		277		277		171		277	

All estimations include a constant term and robust standard errors are reported. ***/**/* reports significance at 1/5/10%.

Sub-industry indicators (at the five-digit level) are included in all regressions (outside the labor quality function).

(1) Cobb Douglas

(2) Cobb Douglas with the number of workers broken down into college and non-college workers

(3) Nonlinear estimation of a multiplicative labor quality augmented Cobb Douglas production function.

(4) Cobb Douglas substituting the number of workers by the firm's monthly wage bill

(5) Nonlinear estimation of an additive labor quality augmented Cobb Douglas production function.

(6) Nonlinear estimation of a multiplicative labor quality augmented Cobb Douglas production function, using the wage bill instead of the number of workers.

(7) Nonlinear estimation of a labor quality augmented Cobb Douglas production function using the Olley and Pakes (1996) estimator to correct for input correlation with true productivity

(8) Cobb Douglas with firm age in logs, firm growth over the last 5 years using the Davis and Haltiwanger (1992) measure. Firm employment growth is calculated using the Davis and Haltiwanger (1992) measure: (xt-xt-5)/((xt+xt-5)/2)

⌖ Stata does not calculate standard errors because of likely collinearity in derivatives at the convergence point. We ran bootstrapped standard errors and they were respectively 0.123 for industry tenure 3 to 5 years and 0.375 for industry tenure 6 to 9 years.

TABLE 4: Labor Quality Augmented Cobb Douglas Production Function
for the Fabricated Metals Industry

Dep. variable: Log Value Added	(1) Number of workers		(2) College/non-college		(3) Detailed human capital measures - multiplicative form		(4) Detailed human capital measures - additive form		(5) Wage bill		(6) Detailed human capital measures - wage bill		(7) Detailed human capital measures with OP - multiplicative		(8) Benchmark - firm age and firm growth	
	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.
Log number of workers (# or wages)	0.818***	0.046			0.829***	0.048	0.829***	0.048	0.827***	0.041	0.799***	0.415	0.747***	0.072	0.818***	0.047
Log college workers			0.146***	0.038												
Log non-college workers			0.663***	0.050												
Log physical capital	0.124***	0.027	0.129***	0.028	0.113***	0.028	0.114***	0.028	0.100***	0.026	0.097***	0.026	0.060*	0.033	0.126***	0.027
Firm age (log)													-0.066**	0.032	-0.018	0.025
DH growth 5 years															-0.003	0.047
Female					0.552*	0.308	1.883	2.311			0.783**	0.351	0.862	0.502		
College & master					1.033	0.681	4.199	5.258			0.041	0.358	1.094	1.012		
Community college					0.354	0.599	1.321	2.864			0.205	0.469	-0.794	0.755		
Vocational					0.232	0.315	0.806	1.578			-0.254	0.175	0.274	0.571		
Experience					0.014	0.014	0.043	0.063			-0.006	0.007	0.018	0.026		
Firm tenure 1 to 2 years					0.427	0.400	1.454	2.158			0.052	[illegible]	0.520	0.860		
Firm tenure 3 to 5 years					0.284	0.381	1.157	1.905			0.141	0.362	0.165	0.835		
Firm tenure 6 to 9 years					-0.128	0.410	-0.965	1.818			2.519***	0.747	-0.046	0.792		
Firm tenure 10 years and up					0.248	0.367	1.384	2.020			-0.099	0.295	0.168	0.874		
Industry tenure 1 to 2 years					0.018	0.224	-0.086	0.778			0.052	0.347	-0.266	0.570		
Industry tenure 3 to 5 years					0.591*	0.346	2.186	2.449			0.141	[illegible]	0.428	0.936		

Industry tenure 6 to 9 years					0.973**	0.439	3.706	3.844			-1.158***	0.099	0.544	1.075		
Industry tenure 10 years and up					-0.177	0.228	-1.083	1.394			-0.099	⌖	-0.323	0.656		
SD (ω)													0.503			
SD (η)													0.948			
Industry dummies	5-digit		5-digit		5-digit		5-digit		5-digit		5-digit		5-digit		5-digit	
R-squared	**0.726**		**0.722**		**0.740**		**0.740**		**0.750**		**0.758**		**0.718**		**0.727**	
# observations	548		548		548		548		548		548		372		548	

All estimations include a constant term and robust standard errors are reported. ***/**/* reports significance at 1/5/10%.

Sub-industry indicators (at the five-digit level) are included in all regressions (outside the labor quality function).

(1) Cobb Douglas

(2) Cobb Douglas with the number of workers broken down into college and non-college workers

(3) Nonlinear estimation of a multiplicative labor quality augmented Cobb Douglas production function.

(4) Cobb Douglas substituting the number of workers by the firm's monthly wage bill

(5) Nonlinear estimation of an additive labor quality augmented Cobb Douglas production function.

(6) Nonlinear estimation of a multiplicative labor quality augmented Cobb Douglas production function, using the wage bill instead of the number of workers.

(7) Nonlinear estimation of a labor quality augmented Cobb Douglas production function using the Olley and Pakes (1996) estimator to correct for input correlation with true productivity

(8) Cobb Douglas with firm age in logs, firm growth over the last 5 years using the Davis and Haltiwanger (1992) measure. Firm employment growth is calculated using the Davis and Haltiwanger (1992) measure: (xt-xt-5)/((xt+xt-5)/2)

⌖ Stata does not calculate standard errors because of likely collinearity in derivatives at the convergence point. We ran bootstrapped standard errors and they were respectively 0.370 for firm tenure 1 to 2 years, 0.434 for industry tenure 3 to 5 years and 0.127 for industry tenure 10 years and up.

TABLE 5: Labor Quality Augmented Cobb Douglas Production Function for the Furniture Industry

Dep. variable: Log Value Added	(1) Number of workers		(2) College/non-college		(3) Detailed human capital measures - multiplicative form		(4) Detailed human capital measures - additive form		(5) Wage bill		(6) Detailed human capital measures - wage bill		(7) Detailed human capital measures with OP - multiplicative		(8) Benchmark - firm age and firm growth	
	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.	coef.	std. err.
Log number of workers (# or wages)	0.867***	0.053			0.886***	0.052			0.894***	0.046	0.881***	0.051	0.776***	0.094	0.872***	0.054
Log college workers			0.256***	0.042												
Log non-college workers			0.603***	0.059												
Log physical capital	0.115***	0.033	0.128***	0.033	0.102***	0.033			0.078***	0.030	0.066**	0.032	0.036	0.038	0.118***	0.034
Firm age (log)													-0.046	0.037	-0.031	0.034
DH growth 5 years															0.063	0.059
Female					0.065	0.262					0.332	0.317	0.451	0.589		
College & master					2.800***	1.053					1.323**	0.615	2.598	2.193		
Community college					1.443	0.934					0.682	0.627	4.361**	1.815		
Vocational					1.034	0.676	Did not converge				0.484	0.421	0.674	0.907		
Experience					0.010	0.018					-0.005	0.009	0.055	0.088		
Firm tenure 1 to 2 years					1.797*	0.923					0.224	0.526	3.917	2.593		
Firm tenure 3 to 5 years					2.216**	0.989					1.778	1.446	3.242*	1.800		
Firm tenure 6 to 9 years					1.406*	0.723					2.427*	1.413	1.568	1.207		
Firm tenure 10 years and up					1.184**	0.595					1.430	1.155	0.941	0.973		
Industry tenure 1 to 2 years					-0.334	0.595					0.224	[illegible]	-0.888***	0.141		
Industry tenure 3 to 5 years					-0.213	0.407					0.560	0.453	-0.892***	0.282		

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
Industry tenure 6 to 9 years					-0.016	0.461					-0.804*	0.429	-0.702*	0.399		
Industry tenure 10 years and up					-0.569*	0.296					-0.755**	0.368	-0.969***	0.354		
SD (ω)													0.429			
SD (η)													1.056			
Industry dummies	5-digit		5-digit		5-digit		5-digit		5-digit		5-digit		5-digit		5-digit	
R-squared	**0.812**		**0.819**		**0.843**		**-**		**0.847**		**0.858**		**0.826**		**0.814**	
# observations	254		254		254		254		254		254		171		254	

All estimations include a constant term and robust standard errors are reported. ***/**/* reports significance at 1/5/10%.

Sub-industry indicators (at the five-digit level) are included in all regressions (outside the labor quality function).

(1) Cobb Douglas

(2) Cobb Douglas with the number of workers broken down into college and non-college workers

(3) Nonlinear estimation of a multiplicative labor quality augmented Cobb Douglas production function.

(4) Cobb Douglas substituting the number of workers by the firm's monthly wage bill

(5) Nonlinear estimation of an additive labor quality augmented Cobb Douglas production function.

(6) Nonlinear estimation of a multiplicative labor quality augmented Cobb Douglas production function, using the wage bill instead of the number of workers.

(7) Nonlinear estimation of a labor quality augmented Cobb Douglas production function using the Olley and Pakes (1996) estimator to correct for input correlation with true productivity

(8) Cobb Douglas with firm age in logs, firm growth over the last 5 years using the Davis and Haltiwanger (1992) measure. Firm employment growth is calculated using the Davis and Haltiwanger (1992) measure: (xt-xt-5)/((xt+xt-5)/2)

⌗ Stata does not calculate standard errors because of likely collinearity in derivatives at the convergence point. We ran bootstrapped standard errors and obtained 0.682 for industry tenure 1 to 2 years.

TABLE 6 - Productivity Dispersion with Labor Quality Controls for Eight Industries

		Food and Beverages				Publishing and Printing				Fabricated Metals				Furniture			
		R^2	sd (log TFP)	sd (logTFP) as % of sd(logVA)	q_90/q_10	R^2	sd (log TFP)	sd (logTFP) as % of sd(logVA)	q_90/q_10	R^2	sd (log TFP)	sd (logTFP) as % of sd(logVA)	q_90/q_10	R^2	sd (log TFP)	sd (logTFP) as % of sd(logVA)	q_90/q_10
COBB DOUGLAS	(1) Number of workers	0.86	0.57	37%	3.48	0.84	0.50	41%	3.43	0.73	0.51	52%	3.70	0.81	0.46	44%	3.34
	(2) College/non-college (numbers of workers)	0.87	0.54	36%	3.31	0.84	0.50	40%	3.68	0.72	0.52	53%	3.76	0.82	0.45	42%	3.30
	(3) Detailed human capital measures - Multiplicative	0.89	0.51	34%	2.98	0.86	0.46	37%	3.11	0.74	0.50	51%	3.55	0.84	0.42	40%	3.00
	(4) Detailed human capital measures - Additive	0.89	0.51	34%	3.12	0.86	0.46	37%	3.07	0.74	0.50	51%	3.58	Did not converge			
	(5) Wage bill	0.89	0.51	34%	3.00	0.87	0.45	36%	3.08	0.75	0.49	50%	3.33	0.85	0.41	39%	2.98
	(6) Detailed human capital measures + Wage bill	0.89	0.49	33%	3.00	0.88	0.44	35%	2.95	0.76	0.48	49%	3.45	0.86	0.40	38%	2.98
	(7) Benchmark: firm employment growth, firm age	0.86	0.56	37%	3.38	0.84	0.50	40%	3.40	0.73	0.51	52%	3.72	0.81	0.46	44%	3.25
TRANSLOG	(1) Number of workers	0.87	0.56	37%	3.33	0.84	0.50	40%	3.24	0.73	0.51	52%	3.65	0.81	0.46	43%	3.36
	(2) College/non-college (numbers of workers)	0.88	0.53	35%	3.08	0.85	0.48	38%	3.19	0.73	0.51	52%	3.72	0.82	0.45	42%	3.28
	(3) Detailed human capital measures - Multiplicative	0.89	0.51	34%	3.12	0.87	0.46	37%	3.05	0.74	0.50	51%	3.49	0.84	0.42	40%	2.99
	(4) Detailed human capital measures - Additive	0.89	0.51	34%	3.14	0.87	0.46	37%	3.03	0.74	0.50	51%	3.54	0.84	0.42	40%	3.08
	(5) Wage bill	0.89	0.51	33%	2.93	0.87	0.45	36%	2.98	0.75	0.49	50%	3.27	0.85	0.41	39%	2.96
	(6) Detailed human capital measures + Wage bill	0.89	0.49	32%	2.95	0.88	0.43	35%	2.78	0.76	0.48	49%	3.39	0.86	0.40	38%	2.98
	(7) Benchmark: firm employment growth, firm age	0.87	0.55	36%	3.16	0.84	0.50	40%	3.24	0.73	0.51	52%	3.69	0.81	0.46	43%	3.26
Std.dev. log V.A.		1.52				1.24				0.98				1.06			
# observations (firms)		256				277				548				254			

The unit of observation is a firm in 2001 in the listed industries

The measure q_90/q_10 is the ratio of the 90th to 10th unlogged TFP quantiles

Sub-industry indicators (at the five-digit level) are included in all regressions (outside the labor quality function) except for advertising, which has only one five-digit industry.

Firm age is the log of firm age from company records.

Firm employment growth is calculated using the Davis and Haltiwanger (1992) measure: $(x_t-x_{t-5})/((x_t+x_{t-5})/2)$

TABLE 6 - Productivity Dispersion with Labor Quality Controls for Eight Industries (Continued)

		Machinery and Equipment				Hotels and Restaurants				Building of Complete Construction				Advertising			
		R^2	sd (log TFP)	sd (logTFP) as % of sd(logVA)	q_90/q_10	R^2	sd (log TFP)	sd (logTFP) as % of sd(logVA)	q_90/q_10	R^2	sd (log TFP)	sd (logTFP) as % of sd(logVA)	q_90/q_10	R^2	sd (log TFP)	sd (logTFP) as % of sd(logVA)	q_90/q_10
COBB DOUGLAS	(1) Number of workers	0.81	0.50	44%	3.43	0.73	0.57	52%	3.86	0.58	0.64	64%	5.08	0.68	0.53	57%	3.49
	(2) College/non-college (numbers of workers)	0.81	0.50	44%	3.48	0.72	0.58	53%	3.88	0.58	0.64	64%	5.31	0.70	0.51	55%	3.41
	(3) Detailed human capital measures - Multiplicative	0.82	0.49	43%	3.29	0.79	0.51	46%	3.41	0.61	0.62	62%	4.84	0.79	0.43	46%	2.54
	(4) Detailed human capital measures - Additive	0.82	0.49	43%	3.36	0.79	0.51	46%	3.36	0.61	0.62	62%	4.91	0.79	0.43	46%	2.55
	(5) Wage bill	0.83	0.48	42%	3.19	0.81	0.48	44%	3.17	0.62	0.61	62%	4.70	0.81	0.41	44%	2.20
	(6) Detailed human capital measures + Wage bill	0.83	0.47	41%	3.10	0.81	0.48	43%	3.13	0.63	0.60	61%	4.70	0.82	0.40	43%	2.20
	(7) Benchmark: firm employment growth, firm age	0.81	0.50	44%	3.48	0.73	0.57	52%	3.88	0.58	0.64	64%	5.13	0.68	0.53	57%	3.30
TRANSLOG	(1) Number of workers	0.81	0.49	43%	3.26	0.74	0.56	51%	3.66	0.58	0.64	64%	5.19	0.70	0.51	55%	3.43
	(2) College/non-college (numbers of workers)	0.82	0.49	43%	3.21	0.73	0.57	52%	3.73	0.60	0.62	63%	5.02	0.74	0.48	51%	3.10
	(3) Detailed human capital measures - Multiplicative	0.82	0.48	42%	3.11	0.80	0.50	45%	3.21	0.61	0.62	62%	4.86	0.80	0.42	45%	2.43
	(4) Detailed human capital measures - Additive	0.82	0.48	42%	3.09	0.79	0.50	45%	3.13	0.61	0.62	62%	4.89	0.80	0.41	45%	2.54
	(5) Wage bill	0.83	0.47	41%	3.14	0.82	0.47	43%	3.08	0.62	0.61	62%	4.76	0.82	0.40	43%	2.15
	(6) Detailed human capital measures + Wage bill	0.83	0.47	41%	3.02	0.82	0.46	42%	2.92	0.63	0.60	61%	4.60	0.83	0.39	42%	2.24
	(7) Benchmark: firm employment growth, firm age	0.81	0.49	43%	3.33	0.74	0.56	51%	3.68	0.59	0.64	64%	5.07	0.71	0.50	54%	3.19
Std.dev. log V.A.		1.14				1.10				0.99				0.93			
# observations (firms)		631				369				535				182			

The unit of observation is a firm in 2001 in the listed industries

The measure q_90/q_10 is the ratio of the 90th to 10th unlogged TFP quantiles

Sub-industry indicators (at the five-digit level) are included in all regressions (outside the labor quality function) except for advertising, which has only one five-digit industry.

Firm age is the log of firm age from company records.

Firm employment growth is calculated using the Davis and Haltiwanger (1992) measure: $(x_t-x_{t-5})/((x_t+x_{t-5})/2)$