VERY PRELIMINARY

Talent Sorting and Skill Complementarity Among Software Engineers

Fredrik Andersson, Cornell University,
Matthew Freedman, University of Maryland,
John Haltiwanger, University of Maryland and NBER,
Paul Oyer, Stanford University and NBER
Kathryn Shaw, Stanford University and NBER

December 2006

Abstract

Conventional wisdom in Silicon Valley suggests that firms need to have some top talent on hand in order to recruit additional top talent. This may be because skilled engineers and other professionals want to work with other top performers and/or because firms can gain from complementarities in production among these employees. We analyze this notion empirically using longitudinal matched employer-employee data on software firms. After decomposing wages into an employee "talent" measure and a firm "quality" measure, we find strong relationships between existing talent, newly hired talent, and firm quality. Firms whose workers exhibit greater talent are able to hire more workers with greater talent. Firms that pay well, controlling for the talent of their workers, are also able to attract relatively talented workers. This suggests that, at least within the software industry, firms are largely segregated according to the skill of their employees.

This document reports the results of research and analysis undertaken in part by the U.S. Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This document is released to inform interested parties of ongoing research and to encourage discussion of work in progress. This research is a part of the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Program (LEHD), which is partially supported by the National Science Foundation Grant SES-9978093 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging Grant 5 R01 AG018854-02, and the Alfred P. Sloan Foundation. The views expressed herein are attributable only to the author(s) and do not represent the views of the U.S. Census Bureau, its program sponsors or data providers. Some or all of the data used in this paper are confidential data from the LEHD Program. The U.S. Census Bureau is preparing to support external researchers' use of these data; please contact U.S. Census Bureau, LEHD Program, Demographic Surveys Division, FOB 3, Room 2138, 4700 Silver Hill Rd., Suitland, MD 20233, USA.

"I would argue that definitely they have the best talent. They invest so much because the more great talent you have, the easier it is to attract even more great talent." -- Joe Kraus, co-founder of Excite Inc. and JotSpot, discussing Google (as quoted in "How Google Woos the Best and the Brightest" by Verne Kopytoff, San Francisco Chronicle, December, 18, 2005.)

I. Introduction

As the quote above suggests, and as we write this, Google's recruiting is like a snowball rolling downhill. The company has earned a reputation for having some of the most talented engineers and business people in Silicon Valley. This reputation then becomes a self-fulfilling prophecy as talented people line up to apply for jobs at Google. Conventional wisdom in Silicon Valley suggests that talent begets talent. In this paper, we analyze empirical proxies of talent among software industry employees and study the degree to which a talented existing stock of workers does, in fact, lead to talented new recruits.

As a theoretical matter, it is not clear whether talent should be evenly distributed across firms or whether top talent should congregate at an elite set of firms. Some firms may find it most efficient to apply the concept of comparative advantage rigorously throughout their organizations. Letting a few highly skilled people lead a group that is representative of the industry or economy in which it operates may allow for the best generation of scale economies of talent exploitation. That is, firms may hire a few stars and then leverage their talent by hiring others who are not star talent. On the other hand, if two people's skills are complements in production, it may make sense to hire a group of highly skilled people together to maximize spillovers in production.

We test this idea of complementarity of worker talent, focusing on the software industry. We use a longitudinal matched employer-employee dataset generated by the Census Bureau which matches Census data with unemployment insurance data from two states. This provides a detailed panel dataset of wages and turnover at over 1,000 software firms. We use the methods developed by Abowd, Kramarz, and Margolis (1999) to decompose wages into a firm effect (which we call firm "quality") and an individual fixed effect (which is our proxy for an individual's "talent.")

Our empirical analysis investigates how the talent of a firm's existing stock of workers affects the talent of its new hires. We ask if firms with higher levels of talent are able to recruit more talent. The panel nature of our dataset allows us to look at the relationship between year-to-year variations in a firm's stock of talent and year-to-year variations in a firm's flow of new talent. We find a very strong economic and statistical relationship between existing and new talent. Firms that have more talented workers recruit more talented workers. We find that a firm with a 10% higher talent level than another firm will recruit workers that are about 5% more talented than the other firm, even controlling for differences in the firms' quality (that is, pay structures.) We also find that higher "quality" firms (that is, those that pay more, controlling for their employees' talent) attract more talented workers. Our estimates suggest that a firm with a 10% higher level of our measure of quality of the firm (or premium pay) than another firm recruits new workers with about 2.5% more talent.

We also look at the hiring of the most talented workers in the distribution. We find that the probability that a new hire will be in the top decile of the talent distribution is strongly related to the average talent level at a firm, the talent level of the 90^{th} percentile person in the firm, and the quality of the firm. The most talented workers go to the "best" firms and those with the most talent.

Our findings suggest that, at least within the software industry, there are important complementarities among the most talented workers. As we discuss below, existing empirical evidence suggests that there are some contexts in which diversity in skills within an organization can increase productivity. However, the overall trend in the economy appears to be segregation of firms by talent. The combination of our results and patterns in the data found by others (particularly Kremer and Maskin, 1996) is consistent with the notion that high skill industries are becoming more important in the economy as a whole and that talent is complementary in these types of businesses.

The next section of the paper provides background on prior studies of co-worker interactions and complementarity. Section III describes our data and the empirical methodology we use to analyze talent. Section IV contains our core empirical analysis of the relationship between existing talent, talent of new hires, and firm quality. We conclude and discuss future plans in Section V.

II. Background

A series of recent papers has presented abundant evidence that productivity in the workplace is affected by who works with whom. For example, Mas and Moretti (2006) study supermarket workers and show that a given person's productivity goes up when she is matched with higher productivity co-workers. Falk and Ichino (2006), who perform an experiment involving a simple task, find that workers are more productive when working with others than when working alone. Hamilton, Nickerson, and Owan (2003) study the implementation of team-based systems at a garment factory. They find that team-based productivity is generally higher than productivity under an individual piece-rate system. They find that within-team skill diversity enhances productivity. While all these studies suggest that teamwork increases productivity, they also all find that productivity is particularly enhanced when highly productive workers are teamed with less productive counterparts.[1]

These studies suggest that total productivity will be highest when low-skill and high-skill employees work side-by-side. However, in some environments, we might expect the skill level of different workers to be complements in the production function. Boning, Ichniowski, and Shaw (2007), for example, show that the benefits of team-based organization in the steel industry are greatest when the production process is particularly complicated. This suggests (though it does not prove) that bringing high skill workers together will maximize productivity.

Other places where we might expect workers' skill levels to be complements are academia and professional service firms. Clients bring routine problems to low price, local law firms and more complex issues to firms with high skill lawyers at all levels. The best students go to universities where the best researchers and teachers work. And, as the quote at the start of the paper suggests, conventional wisdom says that top talent congregates together in the software industry. Why would top "talent" congregate in individual firms? Why would this phenomenon exist in some industries (such as academia and, as we will argue, software) but not others (including supermarket checkers and garment factory workers)?

There are at least two (non-mutually-exclusive) reasons why talent would congregate or, in other words, firms would be segregated by worker skill. First, talented workers may be complements in the production function. For example, suppose that firms have two workers each and that productivity at a given firm is simply the product of the two workers' skill levels. Suppose there are two firms, four workers, and that two workers have skill H and two have skill level L where H > L. Then it is efficient for the two high-skill workers to be at one firm and the two low-skill workers to be at the other firm because $H^2 + L^2 > 2HL$. Naturally, the world is

---

[1] Hayes, Oyer, and Schaefer (2006) show that turnover within groups of managers is correlated, suggesting that workers are more productive with some co-workers than others. Their analysis does not measure how productivity on these management teams is related to within-team talent diversity.

much more complicated than this, but any production function where worker skill is complementary will lead to some degree of skill segregation across firms.

Kremer and Maskin (1996) develop a model that includes this complementary of worker skill. Their goal is to explain some stylized facts suggesting that, as the overall returns to skill have increased substantially in the United States in recent decades, workers have become more segregated by skill within individual firms. That is, though some of the empirical work mentioned above suggests that talent levels should be diverse within firms, the broader evidence suggests that firm talent levels have become more segregated over time. As they write, "Economic activity has shifted from firms such as General Motors, which use both high- and low-skill workers, to firms such as Microsoft and McDonald's, whose workforces are much more homogeneous." Relative to the 1960's and 1970's, the US economy is now more starkly divided between a set of firms dominated by high-skill workers and a set dominated by low-skill workers.

Kremer and Maskin's (1996) model explains this phenomenon by incorporating skill complementarity and two other ingredients. They assume that workers of different skill are imperfect substitutes for one another and that tasks vary in their marginal value of skill. This simple model implies that, if skill becomes *more* variable within the economy, skill becomes *less* variable within firms. This is because, though specialization (that is, assigning high skill workers to high level jobs and low skill workers to other jobs) always increases productivity, the complementarity from combining high skill workers can outweigh this effect when skill variance gets high enough.

The software industry is an "economy" that seems likely to meet the assumptions of this model and to have a high level of variation in skill. Software is typically created in teams. Workers are dependent on one another to create the final output but individual output cannot be measured, limiting the use of individual incentives. Within software development teams, specialized skills are required for various parts of the project. As a result, workers are imperfect substitutes. The value of skill in these different tasks and the value of skill across firms vary considerably (see Andersson, et al, 2006).

Garicano and Hubbard (2006) also emphasize the value of worker sorting among knowledge workers. They show that among lawyers, firms that pay high wages to senior partners are also likely to pay high wages to associates. High talent partners appear to leverage their talent more by working with high talent associates. However, Garicano and Hubbard (2006) go a step further and show that it occurs through hierarchies, as highly paid partners are more likely to work with larger numbers of associates. In our work below, we do not address the issue of talent magnification, but we note that software firms may be less hierarchical.

While worker complementarity and the Kremer and Maskin (1996) model provide one explanation for why talented workers congregate, the quote above suggests an alternative reason why talented workers may congregate. That is, workers may simply prefer to work with talented co-workers. Talented workers can expect higher compensation than other workers. They may be able to take this higher compensation in the form of traditional monetary rewards. But they may also consume some of the value of their human capital by working with each other. That is, the quote suggests that people value working with talented colleagues. But talented people may be willing to "pay" more (in terms of a compensating wage differential) for the privilege of working with smart people than other workers are because their skill gives them a greater ability to afford this luxury. In this case, even if productivity is not affected by how talent combines, it is efficient to combine talented workers with one another.[2]

At this point, we will not attempt to determine whether worker complementarity comes from the production function or from employee preferences. But the rest of this paper will empirically investigate the degree to which software firms in California and Maryland are segregated by talent. We will measure if, and to what degree, talented workers attract other talented workers. We will also measure whether high paying firms (that is, firms that pay well even adjusting for the quality of their workers) attract talented workers.

III. Data and Empirical Methodology

We use an employer-employee matched data set based on data from the U.S. Census Bureau's Longitudinal Employer-Household Database (LEHD) that integrates information from state unemployment insurance data with Census Bureau economic and demographic data. These data provide longitudinal information on workforce composition at the firm level. We linked these data to employer data for software firms (those with NAICS 5112) from the Economic

[2] See Stern (2004) for evidence that another set of high-skill workers, research scientists, "spend" some of the returns to their skill by accepting lower compensation in order to work in a research-oriented environment.

Census for the Services industry.[3] Currently, we use firms in California and Maryland for the years 1992-2002.

We exploit estimates from the LEHD data infrastructure of each person's "talent", which we will denote θ, using the decomposition methodology developed by Abowd, Kramarz, and Margolis (1999) (AKM hereafter.) Abowd, Lengermann, and McKinney (2003) (ALM hereafter) provide further background and show that these measures capture a much broader array of skills than do traditional measures.[4] The basic idea of this decomposition is to identify person and firm wage effects (which we will call individual "talent" and firm "quality" throughout) based on job switchers over a series of years.[5]

AKM and ALM consider the equation

$$\ln w_{it} = \theta_i + \psi_{\mathrm{J}(i,t)} + x_{it}\beta + \varepsilon_{it} \tag{1}$$

where the dependent variable is the log fulltime, full year wage rate of individual *i* working for employer *j* at time *t* and the function J(*i,t*) indicates the employer *j* of individual *i* at date *t*. θ, which is our measure of talent, is a time-invariant person effect. ψ is a time-invariant firm effect which may reflect any number of factors (e.g., rent sharing, bargaining agreements) that are relevant for the firm. x_{it} is a vector of observable individual characteristics (which includes labor market experience and other controls) and ε is a person/year shock that is orthogonal to all other effects in the model. As AKM and ALM discuss, the person effects and firm effects can be estimated for connected groups. By connected groups, we mean that two workers are connected if there exists a chain of coworkers (potentially a very large chain at many different firms given worker mobility) connecting these two workers. Interestingly, in the U.S. virtually all workers are connected for the estimates reported in this paper.

An individual's portable human capital consists of $\theta + x\beta$. In the language of AKM and ALM, this portable measure of human capital captures all the components of the individual worker that are compensated in the labor market controlling for the firm-specific components of compensation. Note that the talent variable includes factors that are often observable but do not vary over a career (such as gender and education) often observable to the statistician and factors

---

[3] See Andersson, et al, (2006) for more details on the LEHD data for software firms.

[4] In addition, ALM developed a new estimation algorithm that has been refined since then to permit estimation of this decomposition on very large pooled samples. The current estimates are based on pooling all workers and firms in 30 states.

[5] The discussion in this section draws from AKM, ALM and Abowd, Haltiwanger, Lengermann, McKinney, and Sandusky (2006).

that typically are not observable (such as innate talent, work ethic, and school quality.) We will rely heavily on a person's θ to represent their (unidimensional) "talent." ALM and AKM detail the assumptions necessary for interpreting theta as a person's ability.

For each firm, we define several variables in each year, including

- ψ = firm fixed effect
- $\bar{\theta}$ = the average θ for all incumbent workers (workers also employed in prior year)
- θ_{90} = the 90th percentile of firm θ for incumbent workers
- Δ_{9010} = 90-10 inter-percentile range of firm θ for incumbent workers
- Δ_{9050} = 90-50 inter-percentile range of firm θ for incumbent workers
- σ_{θ} = standard deviation of θ within firm for incumbent workers
- $\bar{\theta}_n$ = average θ of new hires at firm in year t
- size = number of incumbent employees at firm (excluding new hires)
- g = Firm's employment growth rate between t-1 and t (log difference)

Table 1 presents year-by-year firm-level averages of $\bar{\theta}$, ψ, Δ_{9010}, Δ_{9050}, size, g, and the unadjusted average log wage. There are about 1,000 firms each year. The number grows steadily throughout the late 1990's and then drops off considerably as the internet bubble bursts and the economy cools. This also probably represents some consolidation around 2000. While the number of firms drops, the average software firm in these two states grows from 35 employees to 72 between 1993 and 2002. Total employment in the sample grows in every year except from 2001 to 2002. The standard deviation of the mean firm size is quite large and it also grows considerably. So there are some very large firms in the sample.

The average "ability", as measured by $\bar{\theta}$, at the typical sample firm is 0.422 for all the years. The AKM decomposition is based on log compensation and the average ability for all workers in all states and all years is normalized to zero. This indicates that a typical worker at an average software firm earns 53% more than an average worker in the economy as a whole, adjusting for labor market experience. Software, at least in these two states, attracts relatively talented workers.

Note that the average ability goes down over time. This is likely due to the fact that, as the typical firm grows, it brings on more support staff. If this is driving the trend in ability, it would suggest that skill complementarity drops over time. That is, more mature firms can most efficiently operate with a relatively broad set of abilities among workers.

The firm effects, ψ, are also in logs and normalized to zero for the economy as a whole. So the average level of 0.321 for the sample implies that the typical software firm in these two states pays 38% more, controlling for employee ability, than the average firm in the economy as a whole. This figure goes up slightly over time. This is to be expected given that ψ is increasing in firm size for the economy as a whole and that firm size is increasing in the sample over time.

Table 2 shows correlations among the key firm-level variables. Not surprisingly, the measures of variation are all positively correlated. Also, the measures of average ability (both for existing and new employees) and size are all positively related to the firm effect.

The means and correlations in Tables 1 and 2 suggest that "better" firms survive and grow over the sample period. The numbers are consistent with firms getting bigger, paying more as they grow, and adding more of a hierarchy with relatively more low-level workers (driving firm average ability down). The correlations also show two other features that are prominent in the regression analysis in the next section. Firms with a high talent dispersion in the firm, or high Δ_{9050}, are less likely to hire high quality new workers, $\bar{\theta}_n$. That is, highly talented new hires are more likely to be combined with other talented new hires. Moreover, high growth firms, g, are more likely to hire highly talented people.

IV. Do Talented Workers Congregate?

A. Existing and New Hire Talent

We now more formally analyze how the talent level of a firm is related to the talent level it is able to attract to the firm. We start with regressions of the form:

$$\bar{\theta}_{n\,jt} = \alpha_0 + \alpha_1 X_{jt} + Z_{jt}\beta + D_{jt}\gamma.$$

The dependent variable is the average talent of workers hired by firm *j* in year *t*. X_{jt} measures a firm's incumbent worker human capital (we use different measures such as mean talent, 90th percentile, and 90-50 inter-percentile range). By incumbent workers, we mean workers at the firm in period *t* that were also employed by the firm in period *t-1*. Z_{jt} are firm-specific variables and D_{jt} are time dummies or other interaction terms. Age of firms ranges from 0 to 5, with all firms set to zero in 1992. Firms that are 5 years or older are all assigned an age of 5. For consistency, the regressions that use the age variables as explanatory variables limit the sample to 1996-2002.[6]

Table 3 contains the results of a set of regressions where X_{jt} is the average talent of workers already at the firm at the start of year *t* (that is, $\bar{\theta}$.) As the table shows, no matter what covariates are included, there is a strong statistical and economic relationship between the talent of the existing stock of workers, $\bar{\theta}$, and the average talent of new workers, $\bar{\theta}_n$. The coefficients are centered around 0.5, indicating that a firm with 10% more talented workers (that is, 10% more highly compensated, even correcting for firm effects) relative to another firm can expect its next hire to be 5% more talented than the other firm.

The coefficients on the firm fixed effect variable (ψ) are also highly statistically significant. They indicate that a firm that pays 10% higher wages as measured by the firm fixed effect in the wage regressions (thus controlling for talent) than another firm will attract new workers with about 2.4% more talent. Bigger firms also attract higher talent, which is consistent with the well-known positive relationship between pay and firm size (see Oi and Idson, 1999.) The year and age effects are generally insignificant (as we might expect when controlling for firm size).

Columns (5) and (6) consider extended specifications that include a measure of the spread in the upper tail of the talent distribution of incumbent workers (Δ_{9050}). Column (5) is analogous to column (2) with this additional variable and column (6) is analogous to column (4). The effect of the main variable of interest ($\bar{\theta}$) is not impacted much by the inclusion of this spread variable but the spread variable itself has a large negative impact. Firms with a large spread in the upper tail of the firm's talent distribution are more likely to hire new workers that are of lower quality. Thus, there is some tendency for firm's with a large spread to reinforce that spread with hiring workers of lower quality (holding the average quality constant). Moreover, highly talented workers go to firms that have a large number of existing highly talented workers, as reflected in the high mean talent of the firm, and the lack of talent dispersion within the firm.[7]

[6] The results without age use the full sample. We have checked the robustness of the results without age to the restricted sample that starts in 1996.

[7] The positive correlation between $\bar{\theta}_n$ and θ_{9050} may well reflect the teamwork and "flatness" of firms. Software developers may work in teams of highly talented programmers or developers. This interpretation differs from that of Garicano and Hubbard (2006) who use data on lawyers to emphasize the value of hierarchy in leveraging the talent of the top members of the firm.

Overall, these results suggest that highly talented workers are more likely to be hired by firms that already have a high mean talent level in the firm. In addition, the highly talented go to firms that have a more homogeneous set of highly talented workers. Growing firms are also more likely to hire talented new employees.

Table 4 runs similar regressions, but uses the ninetieth percentile of the firm's current talent distribution as the measure of firm talent. This does not have a substantial effect on the interpretation of the relationships between pre-existing talent, newly hired talent, and the firm's quality. The firm's talent and quality both have a large effect on the talent of new hires. The effect of current talent on new talent is somewhat smaller but still large and highly significant. The extended specification including the upper tail spread variable yields results similar to Table 3 as well.

B. Hiring of "Stars"

We now focus on the hiring of individual "stars". We define a person as a star if her individual talent (θ) is in the top decile of the talent distribution for all software workers in our sample. We run probits of the form:

$$\Pr ob(star = 1 \mid newly\, hired) = \alpha_0 + \alpha_1 X_{jt} + Z_{jt}\beta + D_{jt}\gamma.$$

The sample is limited to new hires in a given year. The explanatory variables are as described in the earlier analyses. Table 5 contains the results of probits where we use average talent ($\bar{\theta}$) already at the firm at the start of year *t* (columns 1 and 2) and the firm's 90th percentile of talent (θ_{90} – columns 3 and 4) as the measures of average talent at the firm. All displayed coefficients are marginal effects on the probability of a new hire being a star.

Columns 1 and 2 show a strong positive relationship between the average talent at a firm and the likelihood that its new hires will be stars. An increase of 10% in the talent distribution is associated with a 5% increase in the likelihood of hiring a star. Given that only 10% of employees are stars, this is a very large effect. There is also a positive and significant, though not as large, relationship between firm quality and the likelihood of hiring a star. Note, however, that larger firms are *less* likely to have a new hire that is a star. So, while large firms attract greater talent, on average, the new hires at a large firm are less likely to be stars. This is consistent with the most talented people being best utilized starting new companies. It could also be that large firms have already found the leaders they need and don't need to look for new people at the high end of the talent distribution.

Columns 3 and 4 show that these relationships hold when looking at the 90^{th} percentile of the firm's talent distribution. Firms whose 90^{th} percentile worker is 10% higher in the talent distribution are about 7% more likely to hire a star.

C. Alternative Specification

We now rerun the analyses in Tables 3 and 5 using a measure of talent that does not control for as many other factors. Specifically, in Tables 6 and 7, we regress log wages of new hires on the log wages of incumbent workers using analogous specifications to Tables 3 and 5. We regard these results as a form of a robustness check to see whether the same basic patterns hold when using a broader but much cruder measure of human capital – the wage itself. We find the same basic patterns as Tables 3 and 5. Firms that pay high wages to their incumbent workers hire new workers at a high wage. This holds at the mean wage as well as in considering the likelihood of hiring a star worker (defined now in terms of the upper decile of wages).

One interesting aspect of these alternative specifications is that when we also include the firm fixed effect as a control, the findings are almost identical to those in Tables 3 and 5. Comparing, for example, column (4) of Table 3 to column (4) of Table 6, we obtain a coefficient of 0.537 and 0.527 respectively on the mean "human capital" of incumbents (measured as $\bar{\theta}$ in Table 3 and the log(wage) in Table 6). To understand why this is the case, recall that the AKM/ALM specification decomposes log wages into person effects, firm effects, and other controls. Moreover, by construction, the firm effect is the same for incumbents and new hires. Thus, in Tables 6 and 7, when we include the firm effect it is as though we are exploiting only the variation in the person effect and the residual for the incumbents and the new hires. Interestingly, including the variation from the residual in the measure of human capital does not impact the results.

V. Conclusions and Next Steps

A common refrain in the software industry goes, "A workers hire A workers, and B workers hire C workers." That is, firms that have an existing workforce of "A" quality programmers and developers are able to attract additional high quality in their new recruits. Those firms that hire "B" workers are never able to hire the "A" workers (and may attract even less talented workers than they currently have) in part because their firms are not succeeding and in part because workers sort carefully. We find evidence that this is the case: firms with a high existing talent pool attract talented new workers. Moreover, high quality new employees are more likely to join an existing firm that has many higher quality workers. Firms hiring high quality new workers have lower quality dispersion within their firms, suggesting that "A" firms hire a relatively homogeneous set of "A" workers.

Our results suggest that high quality workers are complements in the production of software innovations. Firms that have an existing high-talent pool still find enough value added from new talent to pay the premium wages to attract it. In addition, the results are consistent with highly talented workers preferring to work with other stimulating and talented workers when the output is intellectual property. Either of these things – talent complementarity or dependent preferences – would produce the positive effect of the firm's current talent mix on its ability to hire high quality workers.

Regardless of whether skill congregates due to complementarities in production or due to skill-based preferences, the quote at the start of the paper and our results highlight that, at least in software, starting a new firm with the right team of workers is crucial. Firms that start with high talent are able to build on that talent and hire better workers than firms that do not. Success creates more success. If a young firm is doing poorly, it is very difficult to hire the talent to turn it around. While economists may suggest that firms just need to pay higher wages to attract the talent they need, our data suggests that that may not be sufficient – that wages would have to be very high to attract talent to firms with existing low-talent workers. In future work, we will examine the differences in sorting by firm age and growth more carefully.

Our results also speak indirectly to the topic of rising wage inequality. Recent work shows that wages are becoming more polarized: that the gap between wages at the 90^{th} percentile and the 50^{th} percentile continues to grow every year (Autor, Katz, and Kearney, 2005, 2006). We do not examine the time series change in inequality within software. However, the software industry is likely to contribute to this rising wage polarization. There is extreme sorting in which high talent workers go to high paying firms.

References

Abowd, John M., John Haltiwanger, Julia Lane, Kevin L. McKinney, and Kristin Sandusky, 2006, "Technology and Skill: An Analysis of Within and Between Firm Differences," working paper, November.

Abowd, John M., Francis Kramarz, and David N. Margolis, 1999, "High Wage Workers and High Wage Firms," *Econometrica*, 67(2), 251-333.

Abowd, John M., Paul Lengermann, and Kevin L. McKinney, 2003, "The Measurement of Human Capital in the U.S. Economy," US Census Bureau Technical Working Paper #TP-2002-09.

Andersson, Fredrik, Matthew Freedman, John C. Haltiwanger, Julia Lane, and Kathryn Shaw, 2006, "Reaching for the Stars: Who Pays for Talent in Innovative Industries?" NBER Working Paper #12,435.

Autor, David, Lawrence Katz, and Melissa Kearny, 2005, "Trends in U.S. Wage Inequality: Re-Assessing the Revisionists", NBER Working Paper No. #11,627.

Autor, David, Lawrence Katz, and Melissa Kearny, 2006, "The Polarization of the U.S. Labor Market," *American Economic Review*, 96(2), 189-194.

Boning, Brent, Casey Ichniowski, and Kathryn Shaw, 2007. "Opportunity Counts: Teams and the Effectiveness of Production Incentives," *Journal of Labor Economics*, forthcoming.

Falk, Armin and Andrea Ichino, 2006, "Clean Evidence on Peer Effects," *Journal of Labor Economics*, 24(1), 39-57.

Garicano, Luis and Thomas Hubbard, "The Return to Knowledge Hierarchies," working paper, July 2006.

Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan, 2003. "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, 111(3), 465-497.

Hayes, Rachel M., Paul Oyer, and Scott Schaefer, 2006. "Co-Worker Complementarity and the Stability of Top Management Teams." *Journal of Law, Economics, and Organizations*, 22, 184-212.

Kremer, Michael and Eric Maskin, 1996, "Wage Inequality and Segregation by Skill," NBER Working Paper #5,718.

Mas, Alexandre and Enrico Moretti, 2006, "Peers at Work," NBER Working Paper #12,508.

Oi, Walter Y. and Todd L. Idson, 1999. "Firm Size and Wages," in *Handbook of Labor Economics*, Vol. III, eds. Orley C. Ashenfelter and David Card. Amsterdam: North Holland.

Stern, Scott, 2004. “Do Scientists Pay to Be Scientists?” *Management Science*, 50(6), 835-853.

Table 1: Means and Standard Deviations of Key Variables

Year	#Obs	$\bar{\theta}$	ψ	Δ_{9010}	Δ_{9050}	Size	g	log(*wage*)
1993	1023	0.463	0.270	1.064	0.507	34.8	0.322	10.677
		(0.374)	(0.310)	(0.578)	(0.326)	(116.1)	(1.174)	(0.631)
1994	1017	0.460	0.256	1.073	0.507	36.2	0.261	10.683
		(0.331)	(0.292)	(0.561)	(0.311)	(123.9)	(1.081)	(0.655)
1995	1120	0.453	0.274	1.023	0.492	36.4	0.359	10.713
		(0.353)	(0.298)	(0.557)	(0.311)	(146.8)	(1.425)	(0.635)
1996	1216	0.450	0.297	0.981	0.481	37.7	0.409	10.724
		(0.343)	(0.300)	(0.561)	(0.328)	(170.9)	(2.816)	(0.668)
1997	1233	0.441	0.305	1.009	0.496	42.5	0.255	10.747
		(0.328)	(0.305)	(0.540)	(0.314)	(214.9)	(1.046)	(0.678)
1998	1231	0.409	0.322	1.004	0.492	45.1	0.330	10.802
		(0.341)	(0.317)	(0.548)	(0.330)	(252.6)	(2.793)	(0.681)
1999	1204	0.399	0.345	1.016	0.498	49.9	0.377	10.862
		(0.328)	(0.320)	(0.541)	(0.322)	(285.0)	(2.316)	(0.698)
2000	1186	0.379	0.369	1.021	0.512	55.2	0.447	10.921
		(0.304)	(0.324)	(0.530)	(0.320)	(289.2)	(3.511)	(0.679)
2001	1104	0.376	0.388	1.053	0.526	68.3	0.059	10.937
		(0.305)	(0.310)	(0.519)	(0.307)	(327.3)	(2.063)	(0.679)
2002	932	0.391	0.385	1.038	0.521	71.7	-0.021	10.920
		(0.319)	(0.305)	(0.520)	(0.312)	(367.4)	(0.643)	(0.638)
All	11265	0.422	0.321	1.026	0.502	47.4	0.288	10.799
		(0.334)	(0.312)	(0.546)	(0.319)	(243.3)	(2.145)	(0.669)

See Appendix 1 for variable definitions. #obs is the number of firms.

Table 2: Correlations

	$\bar{\theta}$	$\bar{\theta}n$	Log(Size)	σ_θ	Δ_{9010}	Δ_{9050}	Δ_{5010}	g	ψ
$\bar{\theta}$	1.000	0.422	-0.014	-0.198	-0.157	-0.145	-0.116	0.108	0.154
$\bar{\theta}_a$		1.000	0.071	-0.127	-0.101	-0.126	-0.044	0.044	0.247
log(Size)			1.000	0.040	0.396	0.352	0.305	-0.156	0.210
σ_θ				1.000	0.868	0.661	0.680	-0.052	-0.112
Δ_{9010}					1.000	0.814	0.840	-0.144	-0.044
Δ_{9050}						1.000	0.368	-0.124	-0.062
Δ_{5010}							1.000	-0.116	-0.013
g								1.000	0.062
ψ									1.000

See Appendix 1 for variable definitions. #obs is 11,265 firms.

Table 3: Regressions of Average Talent of New Hires

Dependent variable = $\overline{\theta}_n$ (firm average human capital of new hires)

	(1)	(2)	(3)	(4)	(5)	(6)
$\overline{\theta}$	0.524	0.487	0.497	0.537	0.475	0.518
	(49.22)**	(45.62)**	(38.08)**	(12.94)**	(44.26)**	(12.53)**
Δ_{9050}					-0.110	-0.110
					(9.26)**	(7.77)**
log(Size)		0.011	0.014	0.014	0.020	0.022
		(4.62)**	(4.83)**	(4.76)**	(7.65)**	(6.88)**
ψ		0.240	0.224	0.223	0.227	0.211
		(20.42)**	(15.88)**	(15.79)**	(19.24)**	(14.89)**
G		-0.000	-0.001	-0.001	-0.001	-0.001
		(0.07)	(0.65)	(0.34)	(0.53)	(0.62)
age_1			0.034	0.106		0.104
age_2			0.000	0.000		0.000
			(.)	(.)		(.)
age_3			0.007	0.012		0.018
			(0.38)	(0.38)		(0.60)
age_4			-0.011	0.008		0.017
			(0.63)	(0.28)		(0.63)
Age_5			-0.046	-0.029		-0.020
			(3.15)**	(1.29)		(0.87)
$age_1 * \overline{\theta}$				-0.141		-0.135
				(2.73)**		(2.63)**
$age_2 * \overline{\theta}$				0.007		0.015
				(0.13)		(0.29)
$age_3 * \overline{\theta}$				0.000		0.000
				(.)		(.)
$age_4 * \overline{\theta}$				-0.031		-0.028
				(0.59)		(0.53)
$age_5 * \overline{\theta}$				-0.025		-0.020
				(0.55)		(0.44)
						(3.70)**
Constant	0.048	-0.019	-0.050	-0.072	0.027	-0.031
	(8.38)**	(1.38)	(2.64)**	(2.97)**	(1.84)	(1.27)
			(1.94)	(3.76)**		
Obs	11188	11186	8036	8036	11186	8036
R^2	0.18	0.22	0.22	0.22	0.22	0.23

Each regression includes controls for year fixed effects. Absolute value of t statistics in parentheses

* significant at 5%; ** significant at 1%

Table 4: Regressions of Average Talent of New Hires
Dependent variable = $\overline{\theta}_n$ (firm average human capital of new hires)

	(1)	(2)	(3)	(4)	(5)	(6)
θ_{90}	0.278	0.262	0.260	0.357	0.457	0.522
	(29.38)**	(27.35)**	(22.33)**	(11.45)**	(42.93)**	(17.30)**
Δ_{9050}					-0.488	-0.492
					(35.13)**	(29.44)**
log(Size)		-0.010	-0.003	-0.003	0.014	0.017
		(3.90)**	(0.90)	(1.01)	(5.47)**	(5.25)**
ψ		0.315	0.299	0.298	0.234	0.219
		(25.70)**	(20.38)**	(20.34)**	(19.74)**	(15.44)**
G		0.005	0.003	0.003	0.000	-0.001
		(3.13)**	(1.55)	(1.61)	(0.20)	(0.29)
age_1			0.050	0.075		0.111
age_2			0.000	-0.045		-0.013
			(.)	(1.27)		(0.38)
age_3			-0.003	-0.021		0.017
			(0.16)	(0.53)		(0.44)
age_4			-0.035	0.022		0.051
			(1.93)	(0.64)		(1.60)
age_5			-0.079	0.000		0.000
			(5.16)**	(.)		(.)
$age_1 * \theta_{90}$				-0.078		-0.103
				(1.82)		(2.53)*
$age_2 * \theta_{90}$				0.000		0.000
				(.)		(.)
$age_3 * \theta_{90}$				-0.029		-0.021
				(0.60)		(0.45)
$age_4 * \theta_{90}$				-0.114		-0.078
				(2.71)**		(1.96)*
$age_5 * \theta_{90}$				-0.137		-0.057
				(3.90)**		(3.55)**
Constant	0.011	-0.023	-0.040	-0.080	-0.001	-0.068
	(1.17)	(1.45)	(1.89)	(3.46)**	(0.05)	(3.12)**
			(2.68)**	(2.27)*		
Obs	11188	11186	8036	8036	11186	8036
R^2	0.07	0.13	0.14	0.14	0.21	0.22

Each regression includes controls for year fixed effects. Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

Table 5: Hiring Stars Probits
An observation is a newly hired employee
Dependent variable = 1 if new hire is in the top 10% of talent distribution

	(1)	(2)	(3)	(4)
$\bar{\theta}$	0.103	0.116		
	(33.66)**	(30.74)**		
θ_{90}			0.068	0.078
			(25.37)**	(24.07)**
log(Size)	-0.002	-0.002	-0.003	-0.003
	(6.47)**	(4.71)**	(10.14)**	(7.23)**
ψ	0.019	0.009	0.046	0.039
	(5.95)**	(2.49)*	(15.45)**	(11.06)**
g	-0.001	-0.001	-0.000	-0.000
	(4.82)**	(5.85)**	(0.91)	(2.06)*
age_1		0.001		0.002
		(0.34)		(0.61)
age_3		-0.005		-0.005
		(1.40)		(1.37)
age_4		-0.013		-0.015
		(4.03)**		(4.78)**
age_5		-0.014		-0.016
		(4.98)**		(5.85)**
Obs	245797	188485	245797	188485

Marginal probabilities are displayed. Each regression includes controls for year fixed effects. Absolute value of z statistics in parentheses
* significant at 5%; ** significant at 1%

Table 6: Regressions of Average Talent of New Hires
Dependent variable = firm's log(average wage) of new hires

	(1)	(2)	(3)	(4)
$\overline{\log(wage)}$	0.743	0.483	0.503	0.527
	(75.39)**	(40.86)**	(36.89)**	(15.42)**
log(Size)		0.001	0.009	0.007
		(0.28)	(1.68)	(1.42)
ψ		0.857	0.793	0.794
		(33.63)**	(26.49)**	(26.51)**
g		-0.003	-0.003	-0.003
		(0.94)	(1.11)	(0.89)
age_1			0.075	0.317
			(2.46)*	(0.72)
age_2			0.000	0.000
			(.)	(.)
age_3			0.025	-0.970
			(0.77)	(2.01)*
age_4			-0.003	-0.951
			(0.09)	(2.22)*
age_5			-0.027	-1.155
			(1.06)	(3.16)**
$age_1 * \overline{\log(wage)}$				-0.114
				(2.62)**
$age_2 * \overline{\log(wage)}$				-0.091
				(2.06)*
$age_3 * \overline{\log(wage)}$				0.000
				(.)
$age_4 * \overline{\log(wage)}$				-0.004
				(0.09)
$age_5 * \overline{\log(wage)}$				0.012
				(0.34)
Constant	2.512	4.956	4.754	5.682
	(23.55)**	(39.88)**	(32.89)**	(17.18)**
Observations	11186	11186	8036	8036
R^2	0.34	0.41	0.42	0.42

Each regression includes controls for year fixed effects. Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

Table 7: Hiring Stars Probits
An observation is a newly hired employee
Dependent variable = 1 if new hire is in the top 10% of talent distribution

	(1)	(2)
$\overline{\log(wage)}$	0.032	0.036
	(15.89)**	(15.13)**
log(Size)	-0.004	-0.006
	(12.01)**	(14.67)**
ψ	0.134	0.120
	(30.02)**	(23.08)**
g	-0.000	-0.001
	(2.92)**	(3.60)**
age_1		-0.010
		(2.83)**
age_2		-0.004
		(1.09)
age_3		-0.008
		(2.21)*
age_5		-0.010
		(3.51)**
Observations	245797	188485

Marginal probabilities are displayed. Each regression includes controls for year fixed effects. Absolute value of z statistics in parentheses
* significant at 5%; ** significant at 1%

Appendix 1: Variable Definitions

θ =	the person-specific effect in a regression of wages on a set of controls, as defined in equation (1) in the text
ψ =	firm fixed effect
$\bar{\theta}$ =	the average θ for all incumbent workers (workers also employed in prior year)
θ_{90} =	the 90th percentile of firm θ for incumbent workers
Δ_{9010} =	90-10 inter-percentile range of firm θ for incumbent workers
Δ_{9050} =	90-50 inter-percentile range of firm θ for incumbent workers
Δ_{5010} =	50-10 inter-percentile range of firm θ for incumbent workers
$\bar{\theta}_a$ =	average θ of new hires at firm in year t
σ_θ =	standard deviation of θ within firm for incumbent workers
size =	number of incumbent employees at firm (excluding new hires)
g =	Firm's employment growth rate between t-1 and t (log difference)
age0 =	1 if firm is less than one year old
age1 =	1 if firm is one year old
age2 =	1 if firm is two years old
age3 =	1 if firm is three years old
age4 =	1 if firm is four years old
age5 =	1 if firm is five or more years old