Reaching for the Stars: Who Pays for Talent in Innovative Industries?¹

Fredrik Andersson, Cornell University, Matthew Freedman, University of Maryland, John Haltiwanger, University of Maryland and NBER, Julia Lane, National Opinion Research Center/University of Chicago Kathryn Shaw, Stanford University and NBER

July 2007

Abstract:
Firms that want to innovate successfully need to hire and motivate highly talented workers. This paper makes a key connection between the potential returns to innovation in terms of new products and the structure of compensation to talented employees. We use linked employer-employee data to show that software firms that operate in product markets with high potential upside gains to innovation (as in video games or Internet firms) have a greater return to hiring ‘stars’ than do other firms that operate in stable markets (like mainframe software). Firms operating in product domains with highly skewed positive returns are shown to pay employees more in up-front starting salaries and to offer much higher compensation growth. Thus, these firms appear to pay for initial skills and also to pay much more for experience or loyalty: star workers who stay with these firms are paid more than in other firms. The large effects on earnings are robust to the inclusion of a wide range of controls for both worker and firm characteristics. We also show that in firms that have actually developed products with high revenues, or that have hit “home runs”, the rewards for star talent are even greater.

JEL Codes: J240, J310, L200, L860

¹ Thanks to George Baker, Tim Bresnahan, Charlie Brown, Ben Campbell, Robert Gibbons, Erica Groshen, Phil Hardiman, Edward Lazear, Alex Mas, Paul Oyer, Julie Wulf and seminar participants at Stanford University, Harvard University, MIT, University of Illinois, UC Berkeley, University of Michigan, University of Toronto, Yale University, Washington State University, UC Santa Barbara, NBER, and the Society of Labor Economics meetings, for their helpful comments. This document reports the results of research and analysis undertaken 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 research and to encourage discussion. 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 Grants SES-9978093 and SES-0427889 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging Grant R01-AG018854-02, and the Alfred P. Sloan Foundation. The views expressed on statistical issues are those of the author(s) and not necessarily those 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 supports external researchers’ use of these data through the Research Data Centers (see www.ces.census.gov). For other questions regarding the data, please contact Jeremy S. Wu, Program Manager, U.S. Census Bureau, LEHD Program, Demographic Surveys Division, FOB 3, Room 2138, 4700 Silver Hill Rd., Suitland, MD 20233, USA. (Jeremy.S.Wu@census.gov http://lehd.dsd.census.gov).
1 Introduction

Innovation occurs because highly talented individuals create new products. Yet innovation is also an inherently risky process, requiring a substantial investment of both time and money with an uncertain payoff. Widespread product adoption can mean an enormous payoff to the firm; failure can lead to large losses. Understanding how firms recruit and motivate talented workers in order to reap large returns in product markets is critical to understanding the innovation process in the U.S economy. In addition, understanding the recruitment and sorting of workers across firms with different product strategies is likely to be important for understanding the observed changes in the structure of earnings in the U.S. economy, especially at the upper tail of the earnings distribution (see, e.g., Autor, Katz and Kearney 2006). This paper advances such understanding by using data that is seldom available on both the product market and the labor market – data on product market payoffs, firm hiring practices, and the returns to workers – in the context of the highly innovative and high technology software industry.

The software industry is particularly appropriate for such a study for a number of reasons. One is that software products are the prototypical poster-child for the U.S. advanced technology industry. Another is that product innovation is very closely tied to the talents of the workforce. Finally, there is substantial variation in product payoffs in different segments of the industry. For example, the video game segment of the industry is characterized by very high stakes product development: some games, such as Grand Theft Auto (Vice City), generate hundreds of millions of dollars in revenues, while many games make much less. Other segments of the industry, such as business applications software, have substantially less risk and lower potential gains. In these segments, once a sufficiently large community has adopted an application, software producers have an installed client base that provides a degree of stability for future product development. This paper exploits each of the features, and particularly the within industry variation in potential product market payoffs, to explore how payoff variation translates into how firms attract, motivate, and retain talented workers.

Three sets of information are needed for the study. The first is the hiring and compensation policies of firms. The second is the differential revenue payoff distribution for different types of software products. The third is the link between the two. We achieve this by
exploiting a rich new longitudinal employer-employee matched data source that permits us to track both firms and workers. On the firm side, we have information about the detailed product mix and revenue outcomes of each business unit, which permits us to measure both the firm’s potential payoffs and its actual performance by software product class. On the worker side, we can measure earnings levels (including the contribution of exercised stock options and bonuses) as well as within-job and between-job earnings growth over time.

The theoretical motivation for the empirical analysis begins with the assumption that innovative firms want workers who are good at designing or selecting new projects. The key insight of our model is that some firms value this talent much more than other firms. The value firms place on this talent depends on the characteristics of the product market. If the firm produces in a product market in which innovation is rarely rewarded, or in which even a great project will generate small returns, such as business applications software, the project payoff distribution has a relatively low variance and firms in the market attach less value to worker talent. On the other hand, in a product market in which the payoff distribution has a high variance, such as the video game sector, firms attach greater value to individual talent, since those that pick projects well can reap relatively large returns. As a result, highly talented individuals should be matched to firms with high potential value added. Based on these ideas, our model predicts that those firms that operate in software product lines with greater variance in potential payoffs will pay higher starting salaries to attract more talent and pay higher wages to retain and motivate talent.

Our empirical analysis first shows that, consistent with the theory, firms operating in software product lines with greater potential upside gains in value added and that we consequently expect to need to attract greater talent, do indeed pay more in starting salaries than other firms. In particular, these firms select talented workers with a history of prior success; workers who have in the past had high wage levels. Second, we show that firms with higher potential upside gains reward workers more for loyalty; that is, we find that those software workers who achieve the highest earnings do so by remaining at firms in software product lines characterized by greater rightward skewness in payoffs. In short, our results are consistent with the view that innovative firms offer skilled individuals (stars) substantial sums of money up front because they are betting on a high-stakes game of producing winning high-payoff products. The results are also consistent with the view that such firms also reward loyalty with performance
pay, which further increases the likelihood that they will win the high stakes product competition.

This work represents an advance on two fronts. First, it advances the connection between product markets and compensation policy in terms of both focus and richness of data. Previous literature that has linked product market strategies and human resource practices of firms has focused on CEO pay, and has, by and large, linked pay-for-performance as a function of firm size or underlying strategies. A few “insider” studies have generally found evidence of a link between strategy and pay within individual firms (Baker and Hubbard 2003; Stern 2004; Wulf 2002, 2005; Garicano and Hubbard 2005), as have those studies that have employed survey data to examine the subject (MacLeod and Parent 1999). Overall, however, although structural models have described how firm technology or markets should shape personnel policies, but rarely have broadly-based empirical studies made the link between product markets and human resource strategy.

Second, this work closes some of the gap between the theoretical and empirical models of incentive contracts and sorting, particularly in the high technology industry. The theoretical models of incentive pay typically state the conditions under which firms optimally adopt that form of pay (Chevalier and Ellison 1999, Garicano and Hubbard 2005, Fallick et al. 2005, Lerner and Wulf 2005). Empirical researchers may, at best, show that some firms succeed and others fail with a given compensation policy, typically because the efforts to document the adoption of different compensation schemes empirically have been stymied by data limitations. In this study, we suggest that incentive pay plans and sorting based on talent are optimal in firms with high potential payoffs in their respective product markets, and we find empirical evidence consistent with these phenomena.

Finally, although on a broader level and beyond the scope of this paper, it is likely that the human resource practices of firms operating in innovative markets help to shape patterns of earnings inequality in the economy. The link that we outline between product market strategies and compensation policy in innovative markets suggests that variations in product market payoffs not only can act to increase the variance of starting salaries across workers, but also can be used to explain widening pay differentials over time. In particular, our findings could potentially shed light on additional relevant factors impacting the polarization of earnings at the
top end of the earnings distribution (Autor, Katz and Kearney 2006), particularly in advanced
technology sectors such as software where the product payoff stakes vary across products.

The paper proceeds as follows. In the next section, we provide some background facts
about the software industry to help motivate our analysis. We outline a brief model linking
product market risk and pay policy in Section 3, and we provide a detailed description of the data
we use to test the predictions of the model in Section 4. In Sections 5 and 6, we present our
empirical specifications and the results from these specifications. We conclude and discuss the
implications of our work in Section 7.

2 Background

In this section, we present a set of basic facts aimed at describing the wage distribution of
workers and the potential payoff distribution facing firms in different product markets in the
software industry. These facts will help to motivate the approach and analysis that follows.

The first set of facts centers on the distribution of earnings among workers in the
software industry. Mean salaries earned by workers in the industry are relatively high, and that
the distribution is quite skewed: a small subset of workers in the industry receives extraordinarily
high compensation. Panel (a) of Table 1 provides summary statistics about the distribution of
income from the 2000 Decennial Census Public-Use Microdata Sample (PUMS) for workers in
all industries as well as for workers specifically in the software industry.²

As panel (a) reveals, the average worker in the software industry earns more than twice
the average of workers in all other industries, regardless of whether the mean or median is used
as the measure of central tendency. It is difficult to draw conclusions about the variance of
earnings for several reasons. One is that the PUMS data does not include the performance
bonuses and stock options so important in the software industry. The second is that the PUMS
earnings data are topcoded. Finally, PUMS data do not identify other measures important to this
study, such as new-hire earnings, or compensation at the end of workers’ tenure.

In order to address these important deficiencies, we present summary statistics in panel
(b) of Table 1 that are derived from new longitudinal employer-employee matched data that
include bonuses and stock options, are not topcoded, and provide information on

² We focus on full-time workers between 21 and 44 years of age. For the purposes of our analysis, we define the
software industry as SIC 7372 (prepackaged software).
Table 1: Summary Earnings Statistics
Workers 21-44

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median*</th>
<th>90th*</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Industries</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Earnings</td>
<td>40,918</td>
<td>31,891</td>
<td>70,160</td>
<td>183,134</td>
</tr>
<tr>
<td>Wage and Salary Income</td>
<td>38,685</td>
<td>31,466</td>
<td>69,097</td>
<td>173,449</td>
</tr>
<tr>
<td><strong>Software Industry (SIC 7372)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Earnings</td>
<td>80,787</td>
<td>63,782</td>
<td>127,563</td>
<td>334,906</td>
</tr>
<tr>
<td>Wage and Salary Income</td>
<td>80,006</td>
<td>63,782</td>
<td>127,563</td>
<td>333,669</td>
</tr>
<tr>
<td><strong>Computer Software Engineers (Census Occupation Code 102) in the Software Industry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Earnings</td>
<td>90,668</td>
<td>70,691</td>
<td>138,193</td>
<td>369,374</td>
</tr>
<tr>
<td>Wage and Salary Income</td>
<td>90,496</td>
<td>70,160</td>
<td>138,193</td>
<td>369,777</td>
</tr>
</tbody>
</table>

| **Software Industry**                                       |        |         |        |        |
| Starting Earnings (Excludes Left-Censored)                  | 69,353 | 59,665  | 108,692| 82,432 |
| Ending Earnings (Censored and Uncensored)**                 | 344,268| 95,508  | 310,644| 2,051,985|
| One-Year Prior Earnings (Censored and Uncensored)**         | 199,172| 86,796  | 220,760| 1,101,658|
| Prior-Spell Ending Earnings***                              | 60,951 | 51,532  | 100,987| 133,153|
| **Top Decile of Workers (by Ending Earnings) in Software Industry** | | | | |
| Starting Earnings (Excludes Left-Censored)                  | 107,660| 80,899  | 184,951| 142,526|
| Ending Earnings (Censored and Uncensored)                   | 2,532,500| 670,993| 6,688,470| 6,064,204|
| One-Year Prior Earnings (Censored and Uncensored)**         | 750,551| 171,642 | 1,338,380| 2,862,843|
| Prior-Spell Ending Earnings***                              | 98,467 | 73,434  | 164,194| 150,428|

* Average within a 10% band around the true percentile. ** Annualized earnings three quarters prior to last observed full quarter. *** Includes only individuals for whom we observe a prior spell in the data.
beginning and end-of-spell earnings. The longitudinal nature of the data permit us to construct and report four different measures of earnings for software workers, including earnings for new hires (measured as annualized earnings at the start of the observed job spell excluding left-censored records), earnings for experienced workers (measured as annualized earnings at the end of the observed spell), earnings of workers one year prior to the end of the observed spell, and earnings of workers in the last quarter of their prior spells with a different employer. There are left and right censoring issues that we address in the standard ways in our econometric analysis below, but even with these limitations, a number of interesting patterns are evident.

One key finding, evident from panel (b) in Table 1, is that ending earnings have a much higher mean and variance than starting earnings. While this finding is consistent with other data sets, we also find that ending earnings are very skewed to the right. The skewness is especially pronounced for the most highly paid workers (the top decile in terms of ending earnings). The median earnings at the end of a job spell exceed $670,000 and the 90th percentile worker earns nearly $6.7 million, compared with a median of only about $81,000 and a 90th percentile of $184,000 at the beginning of a job spell. This suggests that a select group of workers have enormous average within spell earnings growth; while the growth is substantial for the median worker, it is especially true for workers at the 90th percentile. At least a fraction of the high ending earnings could be bonuses or exercised options upon leaving the firm, which makes examining earnings patterns one year prior to the end of spell also potentially relevant. We return to this point later.

These data are from the Longitudinal Employer-Household Dynamics (LEHD) Program at the U.S. Census Bureau. These data are built from employer-filed unemployment insurance (UI) records, which contain data on all earnings, including bonuses and stock options. Because UI data do not contain hours of work or occupation information, we restrict our sample to workers earning at least $50,000 in the software industry. Moreover, we focus on job spells in the software industry that are ongoing in 1997, as this sample of spells is useful for our later analysis exploiting firm level characteristics. The $50,000 threshold is discussed in more detail below, but it is worth noting that in the PUMS data, two-thirds of all software workers and four-fifths of software engineers have total earnings of at least $50,000. Indeed, the mean of total earnings for software engineers in SIC 7372 earning at least $50,000 is $103,881, only slightly higher than the $90,668 reported in Table 1 for workers at all earnings levels. It is also worth noting that the $50,000 represents the worker’s earnings when we last observed him or her in the data; 36% of those earning $50,000 or more when we last observe them have starting salaries less than $50,000. Fortunately, the results in Table 1 (as well our robustness analysis discussed in more detail below) indicate that by using a relatively simple income cutoff, we can identify the software developers and managers in the administrative data. That is, focusing on workers earning more than $50,000 annually in constant 2001 dollars yields workers that are well identified as software developers and managers.
Another key characteristic regarding the nature of compensation in the software industry that can be uncovered using LEHD data is that the pay of software workers rises markedly with tenure. Figure 1 compares the earnings distribution of compensation for new hires (excluding left censored spells) to the distribution for experienced workers (ending earnings if the spell is not censored and the last observed earnings otherwise). While 70% of starting earnings are below $75,000, only 29% of experienced workers earn below $75,000 (experienced workers have an average tenure of five years). Similarly, only 4% of starting earnings are above $150,000, but 21% of experienced workers earn above that amount. Since starting earnings include the earnings paid to new but experienced workers, compensation rises markedly with tenure.

Figure 1: Distribution of Starting Earnings and Experienced Earnings
SIC 7372, Experienced Workers 21-44 Earning $50,000+

A final pertinent feature of the software industry is that there appears to be a high variance to the gains to innovation in a number of product lines. As an illustrative example, we present in Table 2 the distribution of revenues for the top ten video games in 2002. The distribution is highly skewed, even restricting attention to the top ten games. Indeed, the top game earned nearly five times as much as the tenth on the list.
Table 2: Top Video Games in 2002 Ranked by Sales Revenues

<table>
<thead>
<tr>
<th>Game</th>
<th>Firm</th>
<th>2002 Revenues (Millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grand Theft Auto Vice City</td>
<td>Take 2</td>
<td>$218</td>
</tr>
<tr>
<td>Grand Theft Auto 3</td>
<td>Take 2</td>
<td>$120</td>
</tr>
<tr>
<td>Madden NFL 2003</td>
<td>Electronic Arts</td>
<td>$119</td>
</tr>
<tr>
<td>Medal of Honor</td>
<td>Electronic Arts</td>
<td>$73</td>
</tr>
<tr>
<td>Kingdom Hearts</td>
<td>Square Enix</td>
<td>$59</td>
</tr>
<tr>
<td>Spider Man</td>
<td>Activision</td>
<td>$54</td>
</tr>
<tr>
<td>Halo</td>
<td>Microsoft</td>
<td>$51</td>
</tr>
<tr>
<td>SOCOM Seals</td>
<td>Sony</td>
<td>$50</td>
</tr>
<tr>
<td>Super Mario Sunshine</td>
<td>Nintendo</td>
<td>$49</td>
</tr>
<tr>
<td>Tony Hawk</td>
<td>Activision</td>
<td>$46</td>
</tr>
</tbody>
</table>

Data based on Merrill Lynch's "Reinstating Coverage of Video Game Industry" report (January 21, 2004).

In the consumer video game market, the costs of consumers switching to a new game is minimal, and hence firms in the market have enormous potential gains if the product does succeed in the market. However, the same is not true for firms that produce, for instance, enterprise resource software for large mainframe computers. As we show below, their payoff distribution is much less skewed. Before turning to our results, though, the following section outlines a model that links the rightward skewness of firms’ potential payoff distributions to their hiring and compensation policies, and in particular to their propensity to reward talent and loyalty.

3 Model of Innovation

We model the process of producing innovative software products, though this process may well apply to innovations undertaken by most knowledge workers. The fundamental characteristic of software production is the uncertainty that arises because of firms’ inability to predict whether an innovative product will pay off. In software innovation, there are two integral groups of employees, each of which undertakes risky projects. On the one hand, programmers and engineers must begin working on a new software project not knowing whether they will develop a great product. On the other hand, managers must allocate funds to research

---

4 There are other related forms of uncertainty about product market payoffs that fit within the framework of our model. Suppose, for example, that a component of the uncertainty relates to whether workers implement a new idea effectively. In this case, the talented programmers may be those that implement the idea well (e.g., without problematic bugs or other product market features that would have an adverse impact on the returns from the product).
projects not knowing whether the resulting products will succeed in the market. Thus, a model of project selection pertains to the work of programmers and engineers as well as managers.

Given the uncertainty about the likelihood of success for a given project, the key role of an employee seeking to make innovations is to create or pick the best projects. A model by Lazear (2005) demonstrates how employees who are skilled at creating or picking projects should be sorted among firms operating in high variance payoff markets. Assume that projects can have two outcomes, a good outcome that occurs with probability $P$, and a bad outcome that occurs with probability $(1-P)$. Uncertainty derives from the fact that ex-ante, software programmers and managers cannot identify which projects are good and which are bad. As a result, they can make false positive errors, denoted $H'$, by accepting projects that they believe are good but that later turn out to be bad as well as false negative errors, denoted $1-H$, in which they reject a project that would have turned out to be a good project. More specifically, a false positive is defined as

$$H' \equiv Pr(accept\ a\ project \mid project\ is\ actually\ bad)$$

A false negative, on the other hand, is defined as

$$1 - H \equiv 1 - Pr(accept\ a\ project \mid project\ is\ actually\ good)$$

Assume that if a firm chooses to undertake a good project and it pays off, the firm earns $X$. If, on the other hand, a firm chooses to undertake a project that turns out to be bad, it costs the firm $Y$. A firm has zero costs and zero revenue if it rejects projects early. Given these probabilities and net revenues, the expected payoff for a firm is

$$E(payoff) = PHX - (1-P)H'Y + P(1-H)*0 + (1-P)(1-H')*0$$

which simplifies to

$$E(payoff) = PHX - (1-P)H'Y$$
Firms that achieve a high payoff are those that have a high value of \( PHX \). Firms that fail, meanwhile, are those that have a high value of the losses, \((1-P)H'Y\).

We define a “star” worker as someone who provides a high expected value added to the firm. Therefore, in the model of innovation, a “star” worker is an individual who has a higher probability of accepting truly good projects and a lower probability of accepting truly bad projects. This ability could stem from innate talent, be developed as human capital on the job through learning, or arise from higher effort in response to incentives. In any event, star programmers must develop great projects and star managers must allocate resources to them. Both sets of skills are important determinants of success in the software industry. Thus, we define

\[
H + \varepsilon \equiv \text{Star’s Pr(accept or develop a project | project is actually good)}
\]

and

\[
H' - \varepsilon \equiv \text{Star’s Pr(accept or develop a project | project is actually bad)}
\]

where \( \varepsilon \) captures the quality of the star worker, or the talent that person has in picking projects relative to non-star workers. Therefore, the value of selecting a star employee relative to a non-star employee, \( \Delta \), is the incremental expected payoff,

\[
\Delta = [P(H + \varepsilon)X - (1 - P)(H' - \varepsilon)Y] - [PHX - (1 - P)H'Y]
\]

or, more simply,

\[
\Delta = \varepsilon[PX - (1 - P)Y]
\]

Hence, firms in high variance payoff markets value star talent the most, since firms that have either high potential payoffs from good project selection (large $X$) or large potential losses from bad project selection (large $Y$) gain the most in expected value added from having stars with extra talent \( \varepsilon \).
We illustrate this implication with a continuous distribution of payoffs in Figure 2. The continuous distribution of payoffs is consistent with the model above in that, while a firm might have a range of possible projects with different potential payoffs, any given project might have the types of payoffs and probabilities previously described.

The bold line in Figure 2 (a) shows a high variance payoff distribution, while the bold line in Figure 2 (b) shows a low-variance payoff distribution. The dotted lines in (a) and (b) are the changes in the distributions attributable to star talent. In each case, hiring a star would shift the left tail to the right because employing such workers reduces the occurrence of false positives. In other words, for any given project, a star employee reduces by $\varepsilon$ the probability $H'$ of accepting a project that is bad and, as a result, losing $(1-P)Y$. The right tail shifts right because stars also reduce the number of false negatives; that is, for any given project, a star increases by $\varepsilon$ the probability $H$ of accepting a project that is good and has payoff $PX$. The effect of this rightward shift of the payoff distribution is to increase the mean payoff from $PA_1$ to $PA_2$ in Figure 2 (a). This increase represents the gain associated with paying for a star worker and represents a gain that must exceed the cost of hiring that employee.

Figure 2 (b) depicts a narrower underlying project payoff distribution, representing the situation that would occur with less risky projects: those that have both smaller potential gains and losses. When a firm in this kind of product market acquires a star, this low-risk payoff distribution also shifts to the right, but the mean gains are smaller: $PB_2 - PB_1$. As is evident in the figures, the gains to stars are smaller in low-risk product markets than in high-risk product markets, as $(PB_2 - PB_1) < (PA_2 - PA_1)$. In sum, because there are larger gains (or smaller losses) to the selection of great projects in high-variance product markets, stars are more valuable in (a), where potential payoffs are higher, than in (b), where potential payoffs are lower.

**Primary Hypothesis:** Firms operating in product markets that have high variance payoffs should pay higher wages, as high talent workers sort to innovating firms that obtain the highest value added from talent.

A simple summary of this hypothesis is that there is a production function for innovation in which high-priced talent should sort to firms with product markets which have the highest value for such talent. In what follows, we examine whether such sorting occurs, and use high wages as our indicator of “talent”.

12
Figure 2: Shifts in the Payoff Distribution Due to Reductions in False Positive or False Negative Errors

(a) More Risky Payoff Distribution

(b) Less Risky Payoff Distribution
The question of what mechanisms firms with high variance product markets use to attract, retain, and motivate stars is open. We have presented a static model: firms that value project selection more will value talent more and pay high wages for it. We know that the software industry reflects a much more dynamic environment; for example, bonuses and stock options represent a large share of pay when firms want to retain and motivate talent. More generally, firms could devote a lot of resources to selecting star workers carefully, or alternatively they could allocate more resources to training workers on the job and providing strong incentives that reward (and sort) star workers over time as they gain experience within the firm. The simple model above is silent about these mechanisms for obtaining talent. While our data do not provide a breakout of different forms of compensation, our data provide complete earnings histories that include the contribution of stock options and bonuses. The earnings histories permit us to track earnings prior to the current job spell, beginning of spell earnings, and earnings for experienced workers. We use the information on earnings histories in the empirical analysis to help distinguish between alternative ways that firms attract and retain highly talented workers in the software industry.

4 Data

In order to study the connection between the structure of firms’ product market strategies and skill demand, we require a dataset with detailed information on the earnings and employment histories of workers as well as on the product market characteristics of the firms at which these workers are employed. We take advantage of a unique employer-employee matched data set constructed and maintained by the U.S. Census Bureau’s LEHD Program. We further augment the LEHD data with highly detailed firm characteristics from the Economic Census and worker characteristics from the 2000 Decennial Census PUMS.

---

5 See Gibbons, Katz, Lemieux, and Parent (2005) for a model of workers sorting across sectors as firms and workers learn about skills: they show that high wage sectors employ higher skilled workers and also pay higher returns to skills, where skills are measured as combinations of education and experience.
4.1 The Software Industry

We test the hypotheses of our model by focusing on the prepackaged software industry, which corresponds to the four-digit SIC 7372. This narrow focus has a number of key advantages. The first is the payoff structure across different product categories within the software sector. In many software product lines, the payoff distribution is characterized by high variance, which the video game example in Table 2 vividly illustrated, and which we show below in our results.

The second advantage of our focus on SIC 7372 is that firms in the industry are fairly homogenous in their production processes. Software firms in most markets tend focus on the development and sales of R&D intensive innovations. By contrast, many traditional industries, such as automobile manufacturing, have R&D intensive segments, but are not innovative across the board. Thus, in studying software firms, we are studying innovation and the knowledge workers who do it.

A final advantage of studying software is the richness of the available data. In the Economic Census surveys it conducts every five years, the U.S. Census Bureau conducts the collects detailed product line information (described below) as well as information on the size and age of firms that may serve as controls for elements of product market strategies. To match to our individual UI data, we use the 1997 Economic Census. Therefore, we do not have time series panel data on firms.

4.2 The LEHD Data

The LEHD’s longitudinal wage database contains quarterly records of the employment and total earnings of individuals from Unemployment Insurance (UI) data, which is in turn matched to internal administrative records and surveys containing workers’ date of birth, race, and sex. We have complete UI records for ten states for approximately the years 1992 to 2001

---

6 The Census defines SIC 7372 as “establishments primarily engaged in designing and developing prepackaged software, including operating, utility, and applications programs. These establishments may also prepare software documentation for the user, install software for the user, and train the user in the use of the software. Establishments primarily engaged in buying and selling prepackaged software are classified in Wholesale or Retail Trade. Custom computer software services, including computer code authors, are classified in Industry 7371.”

7 We derive data on software workers from a larger database created by the Longitudinal Employer Household Database (LEHD) Program housed at the Census Bureau. We thank Ron Jarmin for sharing information on firm age with the LEHD Program for this project.

8 Because of the sensitivity of these data, they are anonymized before they are used in any Census Bureau projects; all standard identifiers and names are stripped and replaced by a unique “Protected Identification Key.” Only Census
(the precise years vary slightly by state). Key states that employ large numbers of software workers are in the data, and the length of time enables us to construct sufficiently long worker employment and earnings histories to address our research questions.

These data have several important advantages. First, since the scope of the LEHD data is nearly the full universe of employers and workers, we can accurately track the movements of workers through the earnings distribution within firms as well as across firms over time. Second, in contrast to survey-based information, the earnings data represent the earnings that firms actually pay workers as opposed to workers’ memories of their earnings.

A third key benefit of using these administrative data, particularly in the context of this study and the time frame we consider, is that the earnings measures include bonuses and exercised stock options (though not fringe benefits). Obviously, valuing stock options is quite difficult; in this case, the options are valued when they are exercised, or when the employee cashes in the options. We do not have data on when options are granted to employees. However, our sense is that exercised options, rather than options granted to employees, are the preferred measure of pay for our analysis. Indeed, as Oyer and Schaefer (2002) point out, it generally takes about four years for stock options to be fully vested. Further, as Russell (2005) notes, for a typical software company, options are worth nothing for an employee’s first two years, and then are vested at a rate of two percent per month for the remaining three years. Thus, the value of options that a given firm grants depends not only on whether an employee stays with the firm until the options are vested, but also on the growth of the stock price of the company. Given that we are studying software workers, access to options data is a huge benefit.

Bureau employees or individuals who have Special Sworn Status are permitted to work with the data, and there are serious penalties for disclosing the identity of an individual or business. Any research must be for statistical purposes only, and must be reviewed by the Census Bureau and other data custodians. Under Title 13 of the U.S. code, any breach of confidentiality can result in prosecution in which violators are subject to a $250,000 fine and/or 5 years in jail.

There are important exceptions. Most federal employment as well as some agricultural and nonprofit employment is not covered. Independent contractors and self-employed individuals are also not covered. See Stevens (2002) for a full discussion of coverage issues.

To our knowledge, no previous studies have included stock options data for such a wide range of workers across firms. The nature of our data permits us to exploit the fact that in most employment contracts, employees must exercise all options within 90 days of leaving the firm. We are able to track the earnings of employees for those 90 days and we can thus capture the value of all exercised options. For the laws surrounding the reporting of options, see the example from the California Employment Development Department at http://www.edd.ca.gov/taxrep/de231sk.pdf. For an analysis of options granted and data available on option values, see Oyer and Schaeffer (2002).
It is important to emphasize that the LEHD data capture nearly the full universe of covered employers and workers; they are not merely a sample of software workers or firms. Our primary results are based on two datasets, one consisting of 51,589 employment spells and one of 26,276 spells. The samples are based on a number of restrictions aimed at isolating sets of firms and workers well suited to studying the precise connection between product market strategies and compensation policies. First, we limit the data to workers between the ages 21 and 44 in order to model the demand for a fairly homogeneous collection of individuals in the prime of their careers with similar educational vintages. This reduces the sample from the universe of 83,497 spells to 67,452. Second, we limit our individual worker data to those software workers earning more than $50,000 (in 2001 dollars) at the end of their 1997 job spell. Because we are matching individuals to the 1997 Economic Census, we focus on software workers’ spells that span 1997. The rationale behind the $50,000 earnings threshold is that LEHD data do not contain information on hours of work or occupation. Therefore, to limit the data to workers who are likely to be full-time and in more highly skilled occupations, we select those making more than $50,000. We choose the precise threshold based on a close analysis of the distribution of earnings within the relevant set of software occupations (programmers, developers, engineers, and managers) using PUMS data.\textsuperscript{11} Together, the age and earnings restrictions reduce the sample to 51,589 spells.

While most businesses in our sample of workers could be successfully matched to the Economic Census for 1997, a smaller subset had complete information for firms, including size, age, sales, and detailed product line information. There are 26,276 spells for which we have complete information on firm characteristics as well as worker characteristics. All told, 688 unique software firms appear in this sample.\textsuperscript{12}

\textsuperscript{11} The primary occupations on which we focused included Census industry occupation codes 100 (Computer and Information Scientists, Research), 101 (Computer Programmers), and 102 (Computer Software Engineers, Applications and Systems Software), as well as 001-043 (managerial occupations).

\textsuperscript{12} Throughout this paper, when we refer to a firm, we are referring to a firm defined at the State Employer Identification Number (the SEIN, or UI account number), which is the unit of observation in the UI-Wage data. It is an 11-digit number used for reporting taxes at the state level. For single-unit firms, this reflects the entire firm, but for multi-unit firms, the SEIN reflects activity of the firm within a given state. We are able to match the workers to information in to the Economic Censuses since the UI files also include the federal Employer Identification Number (the EIN is on the ES-202 data that is part of the related administrative data system). The EIN is a nine-digit number assigned by Internal Revenue Service (IRS) and used for federal tax purposes by employers, sole proprietors, corporations, partnerships, non-profit organizations, trusts, estates of decedents, government agencies, certain individuals, and other business entities.
For the purposes of several robustness checks, we also construct a subset of data of employees in high-skilled professions based on occupational information in the 2000 Decennial Census confidential long-form survey records. For this sample, we limit our data to those individuals in the software industry whom we can successfully match to the long-form and whom we can identify as software engineers, developers, or managers (irrespective of earnings). We drop those workers in other occupations within the software industry. Because the Decennial Census is a one in six sample of the population in 2000, this sample consists of only 2,638 workers. We use this dataset to check the robustness of our main findings, but due to its small size and the confidentiality of the data, we refer to the results using this sample largely in footnotes in the empirical analysis.

4.3 Measurement

4.3.1 Earnings Levels and Growth

We focus on four measures of earnings levels in the empirical analysis. Recall that we are looking at employment histories for all workers with their 1997 employer. One earnings measure is beginning-of-spell earnings, which corresponds to a given worker’s total earnings in the first full quarter of employment with the employer (with dollar values at annualized 2001 dollars). Beginning-of-spell earnings include the earnings of new hires as well as the earnings of left-censored job spells in our data. The next measure is end-of-spell earnings, which represents a worker’s last full quarter of real annualized earnings with the 1997 employer. Our end-of-spell measure captures the earnings of workers leaving the firm as well as right-censored job spells, and it potentially contains exercised stock options. When an employee leaves a firm, he must exercise his stock options within 90 days, so exercised options rise dramatically at the end of employment. Another measure, which is less likely to include exercised stock options, is earnings one year prior to the end of the spell (or starting earnings if the spell was less than one year long). Finally, we are also interested in the wage histories of workers prior to their 1997

---

13 Throughout the analysis, we use full-quarter earnings, which represent earnings for workers who have been employed by the same employer for an entire quarter; that is, it represents earnings for a worker whom we observe at a firm in quarter t, t-1, and t+1. While this does not rule out part-time work, it does rule out obviously truncated quarters.

14 Sixteen percent of the beginning-of-spell earnings are censored.

15 Forty percent of the end-of-spell earnings are censored.
employer. For those workers for whom we observe a prior employment spell, we measure the level of earnings in the last full quarter of their prior jobs.

When we use the full universe of software workers for the ten years we observe them as opposed to just the subset of workers linked to their 1997 employers, we construct two measures of earnings growth. Earnings growth within the firm, or within-job earnings growth, is the difference between end-of-spell and beginning-of-spell earnings. Between-job earnings growth is the difference between earnings in the first full quarter of a given worker’s new software job and the last full quarter of his or her prior job.

4.3.2 Product Market Payoff Dispersion

Testing the main implications of our model requires estimates of the variance of the expected payoffs of projects in the software product market(s) in which each firm operates. For the prepackaged software industry, the 1997 Economic Census delineates 30 detailed product lines ranging from consumer game and entertainment software to business graphics design and layout software to vertical industry banking software to mainframe computer applications. Software firms in the Economic Census are asked to provide data on their revenue for each of the 30 product lines, and we exploit this information in order to construct a firm-specific measure that reflects the variance of payoffs in each product category. Even though we have only 30 product-class dispersion measures, each firm has a different payoff dispersion that reflects the shares of its revenue coming from these 30 lines.

We create each firm’s product payoff dispersion measure in two steps. First, for each of the 30 product classes, we calculate the 90-50 difference of the log of revenue per worker across all firms in SIC 7372 in the U.S. economy. Because some of these firms produce and sell multiple software products, we treat each product within each firm as though it were a separate revenue stream. Second, for each firm, we calculate payoff dispersion in the product markets in which it operates by weighting the product-specific 90-50 differences for the 30 products by the share of revenue that the firm derives from each product class.

---

16 More specifically, within-job earnings growth is defined as log annualized end-of-spell earnings less log annualized beginning-of-spell earnings, divided by the number of full quarters that a worker was on the job.
17 More specifically, between-job earnings growth is defined as log annualized beginning-of-spell earnings in the new job less log annualized end-of-spell earnings in the old job, divided by the number of full quarters between jobs. Clearly, between-job earnings growth is only defined for those individuals in the sample for whom we observe them in a job prior to their software job (i.e., those whose software jobs are not left censored and those who are not recent entrants or re-entrants into the labor market).
Values of the product revenue dispersion measure for the product lines with the greatest and least dispersion appear in Table 3. The results suggest that there is substantial variation in the skewness of revenues across product lines, implying a high degree of heterogeneity in the degree of risk firms face in product markets even within the narrowly defined industry of software. Further, observed patterns of dispersion across different product lines are in line with expectations, with categories such as video games topping the list of product lines with high payoff dispersion and database and distribution software falling near the bottom.

Table 3: Software Industry Product Line Revenue Dispersion
SIC 7372

<table>
<thead>
<tr>
<th>Product Line Code</th>
<th>Product Line Description</th>
<th>90-50 Difference of Product Line Log Revenue per Worker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1122</td>
<td>Game and Entertainment Software</td>
<td>1.31</td>
</tr>
<tr>
<td>1183</td>
<td>Networking Software</td>
<td>1.17</td>
</tr>
<tr>
<td>1123</td>
<td>Home Productivity Software</td>
<td>1.03</td>
</tr>
<tr>
<td>1161</td>
<td>Banking and Finance Software</td>
<td>0.66</td>
</tr>
<tr>
<td>1142</td>
<td>Distribution Software</td>
<td>0.57</td>
</tr>
<tr>
<td>1184</td>
<td>Database Software</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Based on a 1997 Economic Census data for a national sample of firms.

There are several features of the firm-specific product payoff dispersion measure worth emphasizing. First, this variable reflects each firm’s actual product mix, not its actual revenue. That is, the payoff measure reflects the skewness of revenue per worker in the software product lines in which the firm operates as opposed to its actual revenues per worker. Thus, a firm with a high product payoff dispersion measure is not necessarily a high or low performing business, but rather has a product mix in lines with a more highly skewed distribution of potential payoffs.

Importantly, we use the 90-50 difference as the measure of revenue dispersion in a given product line. While our theoretical model refers to the variance of the entire distribution (the lower tail as well as the upper tail), we focus on the upper tail because we do not observe firms’ losses – revenues are truncated at zero – or firms going out of business. The lack of complete data on the lower tail, then, motivates this focus on the 90-50 difference. Furthermore, although our simple model focuses on both the value of avoiding losses and creating gains, in the software
industry it is reasonable to argue that the key objective of the firm is to produce big wins, since downside losses are typically much smaller (a few years’ salaries for workers) than the upside gains when a product really takes off in the market.18

5 Empirical Approach

The primary hypothesis from the theory is that firms operating in product markets that have high variance payoffs should pay higher wages, as high talent workers sort to innovating firms that obtain the highest value added from talent. We test this primary hypothesis by estimating alternative variants of the following wage equation:

\[
\ln(w_{ij}) = X_i \beta + Z_j \delta + \alpha \sigma_j + \epsilon_{ij}
\]

where \(w_{ij}\) is the real wage for worker i at firm j (measured at some observed point of the employment spell), \(X_i\) is a vector of worker controls, \(Z_j\) is a vector of firm controls, and \(\epsilon_{ij}\) is an error term. The primary hypothesis is that \(\alpha > 0\), because firms with high \(\sigma_j\), the product payoff dispersion measure described above, operate in product markets that pay higher wages for higher talent.19 We describe the worker and firm controls below.

Equation (1) is a reduced form specification reflecting the labor market equilibrium outcomes associated with the hypothesized sorting of high talent workers across firms. This specification is analogous to other specifications in the literature that have suggested that firms of different types have different skill requirements. For example, an active area of research in the literature has been the role of differences in the adoption of advance technology as a source of differences in demand for skills across firms.20 While related, our approach differs given that we are focusing on variation in earnings outcomes across individuals within a very high skilled group driven by differences in product payoff dispersion across firms.

We consider several variants of estimating (1) by using alternative earnings measures (e.g., starting earnings vs. experienced earnings) and by estimating equation (1) at the mean via

18 See Lazear (1998) and Baron and Kreps (1999) for examples of other industries in which downside losses can be huge because the entire brand value of the firm is lost if workers make mistakes (as in an oil spill or product recall due to customer injuries).

19 Note that \(\sigma_j\) captures the variance of payoffs in a given firm’s product class(es), not a firm’s actual historical variance of success.

20 At the firm level, relevant papers include Doms, Dunne and Troske (1997), Dunne et. al. (2004), Bresnahan, Brynjolfsson, and Hitt (2002), Abowd et. al. (2007) and Van Reenen (1996). At the sector level, relevant papers include Berman, Bound and Griliches (1994) and Machin and Van Reenen (1998).
OLS as well as across the distribution of earnings using quantile regressions. The exploration of quantile regressions helps us explore whether the primary hypothesis holds more strongly for workers with relatively high wages given their characteristics. This might be a reasonable implication of the primary hypothesis since it might be that a video game firm only needs a small number of star programmers.

To help motivate and interpret these alternative specifications, it is useful to interpret (1) as being consistent with the underlying wage-skill-effort relationship given by:

\[
\ln(w_{ij}) = X_i \beta + Z_j \delta + (\beta \sigma^p_j) s_{it} + (\eta \sigma^p_j) e_t s_{it} + \epsilon_{ij}
\]

where \(s_{it}\) is skill at innovating and \(e_t\) is effort, and \(t\) is tenure in our data. We estimate equation (1) rather than equation (2) since both \(s_{it}\) and \(e_t\) are unobserved by the econometrician. In doing so, our estimates of \(\alpha\) reflect the contribution of all factors correlated with \(\sigma^p_j\) holding constant the controls. Using equation (2) to interpret equation (1) helps us make the argument that higher \(\sigma^p_j\) probably also reflects higher skills \(s_{it}\), or higher effort, \(e_t\), in the high payoff firms. While we do not estimate (2), using the different measures of compensation permits us to draw inferences about which of these factors is likely playing a role (e.g., the extent to which firms use different human resource practices to raise skills or effort).

One of specifications for equation (1) is to use starting earnings as the dependent variable. In interpreting the starting earnings specification via equation (2), this permits us to consider how firms acquire skills, \(s_{it}\), for innovating. In the market for new hires, high talent workers should sort to high payoff product dispersion firms in return for higher wages. Thus, \(\beta >0\) in the term \((\beta \sigma^p_j) s_{it}\) in (2) implies that the joint effects of higher skill demand in some software product lines (or high \(\sigma^p_j\)) and higher skills for new hires (or higher \(s_{i0}\) for tenure \(t=0\)) raise starting compensation.\(^21\) There is extensive anecdotal evidence that firms in high payoff dispersion product lines, such as Google or Microsoft, are known to spend resources carefully when selecting talent up front.

\(^21\) Because we lack data on worker characteristics (on education, occupation, or job level), we do not build a detailed model wage determination. See Gibbons, Lemieux, Katz, and Parent (2005) for a model of workers’ wages and sector choice as a function of learning and comparative advantage, and for empirical evidence that different sectors - occupations or industries – pay differently for observed and unobserved skills.
We have a number of alternative specifications for equation (1) that use as the dependent variable measures of experienced earnings reflecting the earnings for workers with positive tenure at the firm (and in such specifications we control for tenure). These specifications for experienced workers permit us to examine the degree to which firms with higher product payoff variance attract and retain talented workers through higher offer pay for experienced workers. In interpreting these specifications via equation (2), these effects may be associated with higher skills, higher effort, or both. Many software firms intentionally tie employees down by granting stock options that vest slowly; to preserve a team through a given product cycle, to reduce costly churning, or to provide incentives for effort. All firms may value teamwork. However, firms in product markets with big upside payoffs may value all skills more, including teamwork. Thus, they will pay higher incentives for experience or loyalty; $\eta > 0$ in the term $(\eta \sigma^p) e_i S_{it}$ implies higher returns for effort and skill, $e_i S_{it}$.

Our model predicts that high talent workers should sort to high payoff dispersion firms. Therefore, pay for low-skilled workers should not less of a function of the firm’s product payoff dispersion, as worker sorting into the lower echelons of the skill and earnings distributions is less relevant in high payoff dispersion firms. Since we use pay as a measure of skill, the impact of payoff variance in (2), or the values of $\beta$ or $\eta$, should be greater at higher percentiles of the earnings distribution relative to lower percentiles, conditioning on other worker and firm characteristics. We test this with quantile analysis.

Finally, we examine the growth rates of base pay. Base pay may grow for numerous reasons. Firms may learn about workers’ unobserved talent for innovation over time as tenure rises in the firm, because workers sort both within firms and across firms to the projects or firms that value innovation the most. Therefore, due to worker-firm matching, returns to tenure may be highest in product lines that value innovative talent the most. Returns to tenure may also be highest in product lines that train on the job to be better innovators through firm-specific skills. In our firms, these skills include the ability to work well with other team members on specific

---

22 Case study evidence suggests that some firms offer such contracts. Russell (2005) finds that a larger percentage of a given workers’ pay is performance-based as the worker’s skill level rises.

23 As another example of this technique, Buchinsky (1994) shows that the returns to education are higher at high wage quantiles, while the returns to experience are lower at high wage quantiles. Also, Hallock et al. (2004) show that among CEOs, the sensitivity of wages to firm performance rises as one moves up the earnings distribution.

24 While learning about ability is explored in detail in Gibbons, for our purposes, we translate learning into higher returns to tenure in the most innovative sectors.
development projects. When teamwork is valued, it may take time for employers to identify and build individual talent that reflects workers’ skills complementarities. Overall, firms in high payoff dispersion markets may invest more heavily in the matching and human capital of their employees in light of the high returns to good project selection. In sum, differencing wage equation (2) for experienced base pay minus starting salary predicts $\beta > 0$ for $\beta \sigma_j^p (s_t - s_{i0})$.

In all of our variants of (1), we interpret the effect of payoff riskiness, $\sigma_j^p$, on wages as reflecting differences in demand for skills. However, payoff riskiness could be correlated with other factors that impact wage variation. For example, it may be that the effect of payoff riskiness reflects a compensating wage differential: higher pay for higher future riskiness of income. To account for such effects, we include a number of controls, which we describe in great detail in the next section.

In sum, aiming to keep the model as parsimonious as possible using the available earnings data, we have several testable hypotheses. First, innovating firms may pay higher starting salaries for talent, as they select high talent workers. Second, innovating firms may pay higher incentive pay in the form of bonuses, stock options, or other incentives. Third, innovating firms may pay higher returns to tenure, as firms and workers learn about talent and sort for it within or across firms, or as workers build human capital.

One or all of these hypotheses may be true. Our model of innovation, while predicting that high project payoff firms will value the talent to innovate more, does not specify the mechanism by which these firms will obtain the talent to innovate. Firms in product lines with big upside payoffs may value all skills more, and will therefore pay more for experience or loyalty, as well as higher starting salaries. The goal of our empirical work is to explore several alternative mechanisms for obtaining or rewarding innovation.

5.1 Subsample and Control Variables

The empirical analysis must encompass the variety of alternative versions and refinements of the benchmark specification (1) described above. There are undoubtedly other factors that may be correlated with the sorting of high talented workers into high potential payoff firms that matters for wages. As previously mentioned, our data include the wage histories for all workers who had job spells ongoing in the software industry during 1997. Worker controls include quadratics of tenure at job, tenure in industry, and age, fully interacted with each other.
and with appropriate left and right censoring dummies when the spells for the 1997 job are left or right censored. It is worth noting that although the set of worker controls is limited, we limit the sample to 21-44 year-old workers earning at least $50,000 in the software publishing industry. Thus, we are largely focusing the empirical analysis on individuals in the relevant educational and occupation categories and the homogeneous sample limits our need for controls.25

We also use as control variables a number of firm characteristics, including a quadratic in (log) employment; dummies for firm age; firm employment growth rate; and a dummy for whether the firm is located in a high density, high education, and industrially diversified county. We also include log revenue per worker and the worker churning rate.26 All firm variables are measured in 1997. These control for a variety of other potential influences on earnings outcomes. For example, size, age, and growth rates have been shown in other literature to be indirectly related with rent sharing, while log revenue per worker is thought to be directly correlated with rent sharing.

6 Empirical Results

The wage regression results reported in Table 4 explore the relationship between software workers’ earnings levels and the positive skewness of the payoffs of the product markets in which firms operate. Again, we measure the product payoff dispersion for each firm as the revenue-weighted 90-50 percentile difference in log revenue per worker across all firms that produce products in the same product lines. We examine the impact of product payoff dispersion on mean earnings for software workers in the OLS regression of column (1). We then estimate quantile regressions in columns (2)-(5).

The product payoff dispersion variable has a very significant positive effect on the earnings of average experienced software workers (using end-of-spell earnings for the 1997 job spell, Table 4, column 1)). Moreover, the rightmost columns (2) through (5) of panel (a) indicate that the impact of the product payoff dispersion measure rises sharply with skill level. In other words, software workers at the upper reaches of the earnings distribution gain the most from working at firms in product markets characterized by greater potential upside gains in value.

---

25 We use the data from the Decennial Census of the population for the observed occupations of software engineers and managers to set the $50,000 minimum earnings cutoff.
26 We measure churning as the worker accession rate plus the separation rate less the absolute value of net employment growth.
added. We summarize these quantile results in subsection 6.5 below, after considering different measures of compensation as the dependent variables in the following subsections.

It is also clear that the firm’s current productivity raises the wages of its employees. The current productivity is the firm’s current actual log revenue per worker (row 2) in 1997. Across all earnings measures (in Tables 4 through 6 below), workers tend to earn more when their employers are currently successful; indeed, pay rises very significantly as a function of log revenue per employee. Additionally, the quantile analysis in panel (a) suggests that experienced high-wage workers are paid disproportionately more when their firms thrive. This should be interpreted as a firm fixed effect; firms that are highly productive in 1997 (when we measure firms’ revenues) pay more to workers in adjacent years as well.

These results show that, even after controlling for the firm’s current revenue success, the firm’s potential upside payoff is a very big determinant of pay. That is, firms in software product lines that demand the skill to innovate pay higher wages for that skill.

6.1 Starting Salaries

Starting salaries are higher for the most highly skilled workers attracted to high-payoff dispersion product lines. The coefficient on product payoff dispersion is positive and significant at the 90th and 95th percentiles: firms in the industry that face greater upside gains in the product market are more actively selecting and hiring talented workers. (columns (4) and (5) of panel (b) in Table 4) The impact of the product payoff dispersion measure is positive, but not significant at the mean (column (1) of panel (b) in Table 4).

---

27 Hallock et al. (2004) point out that “higher ability managers [would have] higher pay for performance incentives than low ability managers” (7) due to the lower cost of effort for high ability managers.

28 Van Reenen (1996) shows that innovating firms in manufacturing share rents from innovation by paying higher wages.

29 In interpreting these results, it is useful again to emphasize that, while the product mix payoff risk measure varies across firms, it is not driven by the realized payoffs of the firm but rather the potential payoff distribution based upon the pool of firms with that product mix. This feature substantially mitigates concerns of contemporaneous endogeneity of the payoff mix measure. This payoff risk measure does reflect a choice by the firm (i.e., the choice of product mix), but this choice is likely made either at the founding of the firm or, at the very least, is made infrequently. After controlling for firm performance, the effects of the product market payoff remain unchanged, which should further reduce concerns about endogeneity. Note also that if we limit our data set to firms operating in one product class (sample size 6,513) to focus on firms with clear cut product design objectives rather than heterogeneous firms, our basic results are unchanged.
Table 4: Earnings Level Regression Results

<table>
<thead>
<tr>
<th></th>
<th>1 (OLS)</th>
<th>2 (10th Percentile)</th>
<th>3 (50th Percentile)</th>
<th>4 (90th Percentile)</th>
<th>5 (95th Percentile)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(a) Dependent Variable: End-of-Spell Earnings - &quot;Experienced Earnings&quot;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product Payoff Dispersion</td>
<td>0.3868</td>
<td>0.0537</td>
<td>0.1203</td>
<td>0.8279</td>
<td>1.0215</td>
</tr>
<tr>
<td></td>
<td>(0.0629)***</td>
<td>(0.0340)***</td>
<td>(0.0397)***</td>
<td>(0.0990)***</td>
<td>(0.1341)***</td>
</tr>
<tr>
<td>Log Revenue per Worker</td>
<td>0.1360</td>
<td>0.0349</td>
<td>0.0996</td>
<td>0.1906</td>
<td>0.2241</td>
</tr>
<tr>
<td></td>
<td>(0.0102)***</td>
<td>(0.0055)***</td>
<td>(0.0064)***</td>
<td>(0.0186)***</td>
<td>(0.0279)***</td>
</tr>
<tr>
<td>OLS R-Squared</td>
<td>0.26</td>
<td>0.054</td>
<td>0.094</td>
<td>0.315</td>
<td>0.375</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
</tr>
<tr>
<td><strong>(b) Dependent Variable: Beginning-of-Spell Earnings - &quot;Starting Salaries&quot;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product Payoff Dispersion</td>
<td>0.0526</td>
<td>-0.1848</td>
<td>-0.0141</td>
<td>0.2129</td>
<td>0.3527</td>
</tr>
<tr>
<td></td>
<td>(0.0331)</td>
<td>(0.0460)***</td>
<td>(0.0331)***</td>
<td>(0.0557)***</td>
<td>(0.0860)***</td>
</tr>
<tr>
<td>Log Revenue per Worker</td>
<td>0.0651</td>
<td>0.0399</td>
<td>0.0603</td>
<td>0.0578</td>
<td>0.0707</td>
</tr>
<tr>
<td></td>
<td>(0.0053)***</td>
<td>(0.0075)***</td>
<td>(0.0053)***</td>
<td>(0.0995)***</td>
<td>(0.0152)***</td>
</tr>
<tr>
<td>OLS R-Squared</td>
<td>0.18</td>
<td>0.102</td>
<td>0.111</td>
<td>0.111</td>
<td>0.119</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
</tr>
<tr>
<td><strong>(c) Dependent Variable: End of Prior Spell Earnings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product Payoff Dispersion</td>
<td>0.1840</td>
<td>0.1145</td>
<td>0.1734</td>
<td>0.1406</td>
<td>0.1702</td>
</tr>
<tr>
<td></td>
<td>(0.0563)***</td>
<td>(0.1149)***</td>
<td>(0.0516)***</td>
<td>(0.0722)*</td>
<td>(0.1072)</td>
</tr>
<tr>
<td>Log Revenue per Worker</td>
<td>0.0520</td>
<td>0.0571</td>
<td>0.0470</td>
<td>0.0342</td>
<td>0.0380</td>
</tr>
<tr>
<td></td>
<td>(0.0103)***</td>
<td>(0.0208)***</td>
<td>(0.0094)***</td>
<td>(0.0123)***</td>
<td>(0.0177)**</td>
</tr>
<tr>
<td>OLS R-Squared</td>
<td>0.28</td>
<td>0.220</td>
<td>0.153</td>
<td>0.137</td>
<td>0.130</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
</tr>
<tr>
<td><strong>(d) Dependent Variable: Beginning-of-Spell Earnings – &quot;Starting Salaries&quot;: Adding End of Prior Spell Earnings as Control Variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product Payoff Dispersion</td>
<td>-0.0460</td>
<td>-0.1082</td>
<td>-0.0648</td>
<td>-0.0647</td>
<td>0.0206</td>
</tr>
<tr>
<td></td>
<td>(0.0289)</td>
<td>(0.0390)***</td>
<td>(0.0267)**</td>
<td>(0.0580)***</td>
<td>(0.0768)</td>
</tr>
<tr>
<td>Log Revenue per Worker</td>
<td>0.0481</td>
<td>0.0596</td>
<td>0.0462</td>
<td>0.0279</td>
<td>0.0278</td>
</tr>
<tr>
<td></td>
<td>(0.0053)***</td>
<td>(0.0073)***</td>
<td>(0.0049)***</td>
<td>(0.0109)**</td>
<td>(0.0150)*</td>
</tr>
<tr>
<td>Prior Spell Earnings</td>
<td>0.3573</td>
<td>0.4856</td>
<td>0.4485</td>
<td>0.2579</td>
<td>0.2289</td>
</tr>
<tr>
<td></td>
<td>(0.0049)***</td>
<td>(0.0066)***</td>
<td>(0.0045)***</td>
<td>(0.0157)***</td>
<td>(0.0229)***</td>
</tr>
<tr>
<td>OLS R-Squared</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Observations</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
</tr>
</tbody>
</table>

Worker controls include quadratics of tenure at job, tenure in industry, and age, fully interacted with each other and with appropriate left and right censoring dummies. Firm controls include a quadratic in (log) firm employment, dummies for firm age (<6 years, 6-10, 11+ years), the net employment growth rate, firm average worker churn, and a dummy for whether the firm is in a high density/high education/industrially diverse county. Controls also include time dummies for quarter of separation and/or quarter of accession as appropriate. All spells are the workers 1997 spell when we observe firm-level data.

Bootstrapped (50 repetitions) robust Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Based on LEHD data for ten states.

Column 1 is an OLS regression. Columns 2-5 are quantile regressions.
<table>
<thead>
<tr>
<th></th>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Product Payoff Dispersion</strong></td>
<td>0.3868</td>
<td>0.0537</td>
<td>0.1203</td>
<td>0.8279</td>
<td>1.0215</td>
</tr>
<tr>
<td></td>
<td>(0.0629)***</td>
<td>(0.0340)</td>
<td>(0.0397)***</td>
<td>(0.0990)***</td>
<td>(0.1341)***</td>
</tr>
<tr>
<td><strong>Log Revenue per Worker</strong></td>
<td>0.1360</td>
<td>0.0349</td>
<td>0.0996</td>
<td>0.1906</td>
<td>0.2241</td>
</tr>
<tr>
<td></td>
<td>(0.0102)***</td>
<td>(0.0055)***</td>
<td>(0.0064)***</td>
<td>(0.0186)***</td>
<td>(0.0279)***</td>
</tr>
<tr>
<td><strong>OLS R-Squared</strong></td>
<td>0.26</td>
<td>0.054</td>
<td>0.094</td>
<td>0.315</td>
<td>0.375</td>
</tr>
<tr>
<td><strong>Number of Observations</strong></td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
</tr>
</tbody>
</table>

(b) Dependent Variable: “Lagged Experienced Earnings” (Lagged One Year from End-of-Spell)

<table>
<thead>
<tr>
<th></th>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Product Payoff Dispersion</strong></td>
<td>0.1312</td>
<td>-0.1674</td>
<td>-0.0759</td>
<td>0.4983</td>
<td>0.7846</td>
</tr>
<tr>
<td></td>
<td>(0.0518)***</td>
<td>(0.0445)***</td>
<td>(0.0349)***</td>
<td>(0.1001)***</td>
<td>(0.1512)***</td>
</tr>
<tr>
<td><strong>Log Revenue per Worker</strong></td>
<td>0.1035</td>
<td>0.0476</td>
<td>0.0805</td>
<td>0.1395</td>
<td>0.1353</td>
</tr>
<tr>
<td></td>
<td>(0.0084)***</td>
<td>(0.0069)***</td>
<td>(0.0057)***</td>
<td>(0.0184)***</td>
<td>(0.0300)***</td>
</tr>
<tr>
<td><strong>OLS R-Squared</strong></td>
<td>0.17</td>
<td>0.075</td>
<td>0.084</td>
<td>0.169</td>
<td>0.224</td>
</tr>
<tr>
<td><strong>Number of Observations</strong></td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
</tr>
</tbody>
</table>

(c) Dependent Variable: “Experienced Salary”

<table>
<thead>
<tr>
<th></th>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Product Payoff Dispersion</strong></td>
<td>0.0551</td>
<td>-0.1251</td>
<td>-0.0795</td>
<td>0.3731</td>
<td>0.5954</td>
</tr>
<tr>
<td></td>
<td>(0.0340)</td>
<td>(0.0343)***</td>
<td>(0.0335)***</td>
<td>(0.0697)***</td>
<td>(0.1059)***</td>
</tr>
<tr>
<td><strong>Log Revenue per Worker</strong></td>
<td>0.0763</td>
<td>0.0434</td>
<td>0.0676</td>
<td>0.0881</td>
<td>0.0827</td>
</tr>
<tr>
<td></td>
<td>(0.0051)***</td>
<td>(0.0064)***</td>
<td>(0.0041)***</td>
<td>(0.0120)***</td>
<td>(0.0167)***</td>
</tr>
<tr>
<td><strong>OLS R-Squared</strong></td>
<td>0.17</td>
<td>0.079</td>
<td>0.098</td>
<td>0.142</td>
<td>0.202</td>
</tr>
<tr>
<td><strong>Number of Observations</strong></td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
</tr>
</tbody>
</table>

(d) Dependent Variable: End of Spell Earnings – “Experienced Earnings”, Adding End of Prior Spell Earnings

<table>
<thead>
<tr>
<th></th>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Product Payoff Dispersion</strong></td>
<td>0.1519</td>
<td>-0.0287</td>
<td>0.0295</td>
<td>0.5709</td>
<td>0.6343</td>
</tr>
<tr>
<td></td>
<td>(0.0552)***</td>
<td>(0.0346)</td>
<td>(0.0404)</td>
<td>(0.1172)***</td>
<td>(0.1561)***</td>
</tr>
<tr>
<td><strong>Log Revenue per Worker</strong></td>
<td>0.0837</td>
<td>0.0203</td>
<td>0.0584</td>
<td>0.1505</td>
<td>0.1604</td>
</tr>
<tr>
<td></td>
<td>(0.0102)***</td>
<td>(0.0063)***</td>
<td>(0.0075)***</td>
<td>(0.0240)***</td>
<td>(0.0367)***</td>
</tr>
<tr>
<td><strong>Prior Spell Earnings</strong></td>
<td>0.2810</td>
<td>0.2493</td>
<td>0.2784</td>
<td>0.2808</td>
<td>0.2929</td>
</tr>
<tr>
<td></td>
<td>(0.0094)***</td>
<td>(0.0045)***</td>
<td>(0.0069)***</td>
<td>(0.0313)***</td>
<td>(0.0472)***</td>
</tr>
<tr>
<td><strong>OLS R-Squared</strong></td>
<td>0.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of Observations</strong></td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
</tr>
</tbody>
</table>

Worker controls include quadratics of tenure at job, tenure in industry, and age, fully interacted with each other and with appropriate left and right censoring dummies. Firm controls include a quadratic in (log) firm employment, dummies for firm age (<6 years, 6-10, 11+ years), the net employment growth rate, firm average worker churn, and a dummy for whether the firm is in a high density/high education/industrially diverse county. Controls also include time dummies for quarter of separation and/or quarter of accession as appropriate. All spells are the workers 1997 spell when we observe firm-level data.

Bootstrapped (50 repetitions) robust Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Based on LEHD data for ten states.

Column 1 is an OLS regression. Columns 2-5 are quantile regressions.
Drilling deeper into our data, we see that high-payoff dispersion product lines appear to be both selecting higher quality workers and achieving better job matches. In panel (c), we use a different dependent variable; we use the worker’s end of spell “earnings on the prior job” that they held (prior to their 1997 job) and we regress it on the current 1997 job’s payoff dispersion. For those workers who had previous jobs in our data, the average worker selected to work in high potential payoff product lines had much higher than average earnings in the previous job (i.e., the OLS coefficient is a very significant .184 on the product payoff dispersion). Going a step further, we replicate the starting salary regression of panel (b), but add the “earnings on the prior job” as a control variable for worker quality (panel (d)). In this regression, the effects of the payoff dispersion on starting salaries disappears for high talent workers: high-payoff product dispersion product lines are selecting the best workers they can find or achieving better job matches in hiring, and thus paying higher than average starting salaries. These results are consistent with extensive industry testimony that describes the software industry’s very careful and deliberate hiring practices, all aimed at identifying the right talent and reflecting the high-commitment work environment of the industry (Hoch et al. 2000).

6.2 Incentive Pay for Experienced Workers

We show above that experienced workers earn much more in software product lines with high potential upside gains to innovation. Do these high earnings for experienced workers arise from higher base pay, higher bonuses, or from careful initial hiring of workers? To consider these different effects, in Table 5 we present results for three different measures of pay for experienced workers. The first panel replicates the earnings regression of panel (a) in Table 4 for comparison purposes, but let us emphasize that this measure of earnings contains very significant incentive pay for software workers. The end-of-spell earnings for experienced workers is the earnings in the last quarter in which we observe them in our data with their 1997 employer. Therefore, for 60% of these workers, this is their last quarter working with this employer; the other 40% are right censored in our data and thus stay with the employer. By law, employees must exercise all stock options within 90 days of quitting, thus these data could contain significant exercised stock options as well as bonuses or severance pay. Our regression results show that these software firms operating in product lines with high product payoff

---

30 As a check on possible sample selection, we estimated the OLS specification using the dependent variable as the prior spell earnings with a selection correction and obtained very similar results.
dispersion (high $\sigma^p$) are paying median workers more and are greatly rewarding highly skilled individuals (above the 90th percentile).

When we turn to measures of earnings that include far less incentive pay, we find that high talent workers continue to earn much more in base salary at firms with high potential payoff product lines. To reduce the emphasis on stock options that is in the end-of-spell earnings measure, we measure earnings lagged one year prior to the end of workers’ spells. As the results in panel (b) of Table 5 reveal, although the point estimates of the impact are smaller than for end-of-spell earnings, the same basic results hold.

Next, we create an earnings measure that is closer to base pay by using the minimum of each worker’s end-of-spell and one-year lagged earnings, which we label “experienced salary” because is less likely to contain bonuses. The regression results also show that high talent workers have high experienced salary in high-payoff dispersion product lines (panel (c), Table 5). In all cases, low talent workers lose more in these product lines.

Note finally that the earnings premium that experienced workers obtain in high potential payoff product lines looks to be in part due to careful hiring of these workers up front. When earnings on the prior job is added as a control in the end-of-spell earnings regression, the coefficients on product payoff dispersion decline in size across the quantile regressions (comparing panel (d) to (a)), suggesting that firms hire selectively and that highly talent workers sort into jobs that are ultimately more rewarding.

6.3 Earnings Growth Within Jobs

We next turn to examining the relationship between the nature of the product markets in which firms operate and the earnings growth workers experience on the job. The results of this section echo the previous results.

Within-job earnings growth rates rise sharply with the product payoff dispersion of firms for nearly all workers, though the impact is greatest for workers at the highest earnings quantiles (panel (a) of Table 6). Within-job earnings growth is the difference between workers’ first observed and last observed earnings with their 1997 employer.

Within-job base salary also rises with the product payoff dispersion measure, but the effect is concentrated among the higher talent quantiles of workers. Quantitatively, salary rises much less than total earnings (with bonuses), implying either that human capital and learning is
rewarded through bonuses in high payoff dispersion product lines, or that human capital grows at the same rate across all software product lines.

6.4  **Loyalty Pays: Comparing Earnings Growth Within and Between Jobs**

Software workers are thought to be ‘job-hoppers’ who move frequently between jobs to maximize earnings. Using our data, we can compare the gains to job hopping across firms operating in different software product lines.

While within-job earnings growth was shown to be higher for workers in high-potential payoff firms, between-job earnings growth is lower for workers moving to high potential payoff firms. Between-job earnings growth is measured as the difference between a given worker’s starting salary on his current job (with his 1997 employer) and his ending compensation on his last job. For the median worker, the effect of product payoff dispersion on between-job earnings growth is actually negative and significant, though at higher earnings quantiles it is insignificant.

Therefore, the results suggest that job-hopping for higher earnings may be an excellent long-term strategy, but it is not the best short-term strategy for wage growth. When software workers change jobs, they must stay with the firm a number of years before their compensation rises. In this sense, loyalty pays, as workers who stay with their employers tend to see stronger earnings gains over time. Immediate jumps from one firm to another do not pay. Moreover, the firms that reward loyalty the most are the very firms that operate in high-risk, and thus high potential payoff, product markets. We cannot assess precisely why loyal workers tend to reap the greatest rewards in firms in high-risk markets, and indeed the differential may arise due to factors ranging from variation across markets in the importance of teamwork, firm-specific human capital accumulation, monitoring costs, intellectual property protection, etc. In any case, the results make clear that loyalty in the software industry pays, and pays disproportionately among firms that face the riskiest product markets. Firms in these dynamic markets, therefore, structure compensation not only to select the most talented workers, but also to ensure firms motivate and retain them.
## Table 6: Earnings Growth Regression Results

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>10th Percentile</td>
<td>50th Percentile</td>
<td>90th Percentile</td>
<td>95th Percentile</td>
</tr>
<tr>
<td>Product Payoff Dispersion</td>
<td>0.0706 (0.0120)*****</td>
<td>-0.0060 (0.0073)</td>
<td>0.0481 (0.0065)*****</td>
<td>0.1837 (0.0191)*****</td>
<td>0.2074 (0.0227)*****</td>
</tr>
<tr>
<td>Log Revenue per Worker</td>
<td>0.0102 (0.0020)*****</td>
<td>-0.0058 (0.0012)*****</td>
<td>0.0037 (0.0011)*****</td>
<td>0.0340 (0.0037)*****</td>
<td>0.0471 (0.0048)*****</td>
</tr>
<tr>
<td>OLS R-Squared</td>
<td>0.16</td>
<td>0.073</td>
<td>0.079</td>
<td>0.230</td>
<td>0.280</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
</tr>
</tbody>
</table>

### (b) Dependent Variable: Within Job “Salary” Growth++

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>10th Percentile</td>
<td>50th Percentile</td>
<td>90th Percentile</td>
<td>95th Percentile</td>
</tr>
<tr>
<td>Product Payoff Dispersion</td>
<td>-0.0025 (0.0131)</td>
<td>-0.0136 (0.0092)</td>
<td>0.0068 (0.0047)</td>
<td>0.0499 (0.0146)*****</td>
<td>0.0705 (0.0371)*</td>
</tr>
<tr>
<td>Log Revenue per Worker</td>
<td>-0.0036 (0.0018)*****</td>
<td>-0.0082 (0.0013)*****</td>
<td>-0.0016 (0.0007)*****</td>
<td>0.0120 (0.0005)*****</td>
<td>0.0124 (0.0045)*****</td>
</tr>
<tr>
<td>OLS R-Squared</td>
<td>0.14</td>
<td>0.121</td>
<td>0.092</td>
<td>0.103</td>
<td>0.127</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
<td>26,276</td>
</tr>
</tbody>
</table>

### (c) Dependent Variable: Between-Job Earnings Growth+++

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>10th Percentile</td>
<td>50th Percentile</td>
<td>90th Percentile</td>
<td>95th Percentile</td>
</tr>
<tr>
<td>Product Payoff Dispersion</td>
<td>-0.2169 (0.0476)*****</td>
<td>-0.2352 (0.0653)*****</td>
<td>-0.1725 (0.0270)*****</td>
<td>-0.2597 (0.0921)*****</td>
<td>-0.2215 (0.1965)</td>
</tr>
<tr>
<td>Log Revenue per Worker</td>
<td>0.0188 (0.0087)*****</td>
<td>0.0367</td>
<td>0.0195</td>
<td>0.0080</td>
<td>-0.0221</td>
</tr>
<tr>
<td>OLS R-Squared</td>
<td>0.09</td>
<td>0.039</td>
<td>0.051</td>
<td>0.094</td>
<td>0.113</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
<td>10,803</td>
</tr>
</tbody>
</table>

Worker controls include quadratics of tenure at job, tenure in industry, and age, fully interacted with each other and with appropriate left and right censoring dummies. Firm controls include a quadratic in (log) firm employment, dummies for firm age (<6 years, 6-10, 11+ years), the net employment growth rate, firm average worker churn and a dummy for whether the firm is in a high density/high education/industrially diverse county. Controls also include time dummies for quarter of separation and/or quarter of accession as appropriate.

Bootstrapped (50 repetitions) robust Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Based on LEHD data for ten states.

Column 1 is an OLS regression. Columns 2-5 are quantile regressions.

++ “Experienced” total compensation minus starting compensation with 1997 employer.

+++ The minimum compensation in the last two years of employment (based on quarterly data annualized) minus starting compensation with 1997 employer.

++++ Starting compensation on current job minus ending earnings on the prior job.
The raw data corroborate the regression-based findings that loyalty pays in software. We see this by calculating the total earnings of workers in the last period in which they are observed in the data and separating them into four categories: those who made over $1 million; those who made between $500,000 and $1 million; those who made between $200,000 and $500,000, and those who made between $50,000 and $200,000. We then calculated the proportion of wage growth for each worker that was derived from within-firm earnings growth (i.e., “loyalist”) and that derived from between-firm earnings growth (i.e., “job-hopper”). Of the approximately 4% of the sample who earned over $1 million in the last period, over 95 percent of their wage growth arose within firms, and less than five percent from movement between firms. By contrast, among software workers in the $50-75,000 range, the final pay is achieved by a combination of changing jobs and by wage growth when they stay within a firm and experience wage increases.

Figure 3

Thus, the striking result from Figure 3 is that even within the software industry, workers earn more from loyalty to their employer. That is, by far the greatest wage gains come not from hopping between employers, but rather from staying with an employer and earning higher pay over time. The figure corroborates our regression results: high wage growth arises for workers in

---

31 Our definition of between-firm wage growth is annualized starting compensation minus the ending compensation at the last firm. The starting compensation does not include options granted, so in some sense we could say that we are underestimating the gains to job-hopping if software workers are moving between firms to achieve higher option grants. Nevertheless, our key point is that options granted are not yet compensation — the individual must stay with the firm four years (typically) to vest the options granted and the options must be “in-the-money” as a result of performance. Thus, even if options are granted with job change, the pay is only realized from within firm pay increases — the person must stay and perform.
high payoff product markets when they stay with their current employer—for these workers, the return to loyalty can be very high.

6.5 **Estimated Conditional Earnings Distributions**

The quantile regression results show that the overall shape of the conditional wage distribution differs in firms with high potential payoffs versus low. To make the results of the quantile regressions vivid, we graph the implied wage distributions in Figure 4. Predicted wages at each point in the wage distribution are calculated for firms that operate in software product lines with the “minimum product market risk” (the solid line) and the “maximum product market risk” (the dashed line) for different measures of earnings.

![Figure 4A: Predicted “Starting Earnings” at Minimum and Maximum Product Market Payoff Dispersion](image)

Experienced Workers 21–44 Earning $50,000+, Other RHS Variables at Means

RHS variables include product payoff dispersion, log revenue per worker, employment, employment squared, firm age dummies, firm employment growth rate, location dummies, and time dummies. Variables set at means include log revenue per worker, employment, employment squared, and firm employment growth rate. Assumes firm age of 6–10 years; that the firm is located in a high density, high education, and industrially diverse county; and that the worker accessed in the second quarter.
Figure 4B: Predicted “Experienced Earnings” at Minimum and Maximum Product Market Payoff Dispersion
Experienced Workers 21–44 Earning $50,000+, Other RHS Variables at Means

RHS variables include product payoff dispersion, log revenue per worker, employment, employment squared, firm age dummies, firm employment growth rate, location dummies, and time dummies. Variables set at means include log revenue per worker, employment, employment squared, and firm employment growth rate. Assumes firm age of 6–10 years; that the firm is located in a high density, high education, and industrially diverse county; and that the worker separated in the second quarter.

Figure 4C: Predicted “Experienced Salary” (Minimum Earnings in Last Two Years of Employment) at Minimum and Maximum Product Market Payoff Dispersion
Experienced Workers 21–44 Earning $50,000+, Other RHS Variables at Means

RHS variables include product payoff dispersion, log revenue per worker, employment, employment squared, firm age dummies, firm employment growth rate, location dummies, and time dummies. Variables set at means include log revenue per worker, employment, employment squared, and firm employment growth rate. Assumes firm age of 6–10 years; that the firm is located in a high density, high education, and industrially diverse county; and that the worker separated in the second quarter.
Figure 4 shows that for high payoff firms, the right tail of the earnings distribution is substantially thicker. The upper tail is thicker for all measures of earnings: higher skilled worker have higher starting salaries (Figure 4A), higher earnings with experience (Figure 4B), and higher base salary with experience (Figure 4C). Holding all other controls fixed and using the product line dispersion statistics from Table 3, workers at the 50\textsuperscript{th} percentile employed at a firm producing a product with the highest payoff risk have experienced earnings that are 9 percent higher than a firm producing a product with the lowest product risk. This differential increases to 63 percent at the 90\textsuperscript{th} percentile and to 77 percent at the 95 percentile.

6.6 Discussion

We show that firms that operate in product markets that have high potential payoffs for innovation will pay high levels of compensation and pay higher within-firm growth rates of pay. Our model argues that the gains to talent in project selection are higher in markets when the stakes are higher. Even after controlling for rent sharing (based on their actual revenue per worker), the firm’s potential upside payoff is critical determinant of pay. These results are robust to using different measures of earnings and, as we find in unreported regressions, to using different specifications with varying sets of control variables. Firms in software product lines that appear to demand the skills of innovating, pay higher wages for these skills.

The very high compensation for experienced workers is consistent with a variety of theories: higher marginal products (as in our model), a tournament reward structure, participation costs, and so forth. Additional ways that firms can reward star performers is by assigning them to the most desirable projects or by furnishing them with time to do their own publishable work. Stern (2004) shows that star scientists “pay” to be in more R&D intensive firms by accepting lower wages early in their careers.

Appendix figures A1 through A3 corroborate these results by plotting the difference between the expected earnings for high and low payoff firms. The standard errors around the coefficients at each quantile are also small, as displayed in Figure A4.

In interpreting Tables 4 through 6 and Figure 4 in terms of the magnitudes it is important to emphasize that the reported effects from the quantile regressions yield the implied effect of the variable in question on the conditional quantile distribution. By the conditional quantile distribution, we mean the distribution of earnings taking into account all of the other explanatory variables including the controls. Thus, the coefficients in Table 4 and Figure 3 should not be interpreted to yield inferences about the impact of variables on the unconditional distribution of earnings. For our purposes, the focus on the conditional distribution of earnings is appropriate since we are interested precisely in the impact of product payoff distribution holding the impact of all other factors constant. For further discussion of these issues, see Buchinsky (1994).

Additional ways that firms can reward star performers is by assigning them to the most desirable projects or by furnishing them with time to do their own publishable work. Stern (2004) shows that star scientists “pay” to be in more R&D intensive firms by accepting lower wages early in their careers.

Also in unreported regressions, we find that the results are robust when we look at the subset of individuals for whom we can identify their occupation using Decennial Census data. While limiting the size of our sample substantially, integrating occupation data from the Decennial Census permits us to exclude workers other than programmers, engineers, and managers in software firms. For workers outside these occupations, such as administrative and sales staff, we might expect the link between project success and compensation to be weak.
in a high-performance team, or improved selection of talented workers over time in the firm. The data do not permit us to distinguish among these alternatives with our data; indeed, Russell (2005) provides very detailed evidence for one software company that suggests that all of these factors may enter the earnings of software workers. In any event, the individual who can create or select the best projects will have more skills and more incentive pay in firms with high product payoff dispersion.

While our results are for the software industry, results from broader data sets corroborate and extend several of our key results. Using PSID data across industries, Lemieux et al. (2007) show that, as performance pay has increased over time by firms, the highly educated high-ability long-hours workers have received both more incentive pay and more base pay. These results are especially true for managers and professionals in their work. As in our data, all forms of pay have gone up for the highest skilled workers, suggesting rising demand for skills. In software, we posit that the skill is innovation, or as in Autor, Katz and Kearney (2005) show, it is the ability to solve complex problems.

A few remaining points help put our results in perspective. First, we are modeling the software industry during the boom period of stock options, and our results certainly reflect this fact. Some end-of-spell earnings gains are so large that they must be due to exercised stock options. Therefore, it is important to point out that when we drop workers’ last year of earnings, and look at earnings in a prior period (when stock options are exercised much less), all our qualitative results remain. The basic point of our paper remains: software firms in gaming or Internet domains were producing the biggest innovations before the market boom of the late 1990s and continue to produce gains after the bust. They are still the innovators and still hire and pay for stars. This is still a model of innovation, not simply a model of returns to stock options.

Second, given our data and knowledge of the software market, it is unlikely that our wage regression results reflect a large compensating differential for risk. If we look only at the coefficient of starting salaries on the variance of product payoffs (panel (b) of Table 4), the positive coefficient in our results could well reflect a positive compensating differential for

---

36 All indications are that the firm in Russell’s (2005) study looks very much like the typical large firm in our data. The median age among workers is 33, and tenure ranges from 2.7 to 3.1 years over 1996 to 1999. About 65% of workers were in research and development and 30% were in management or administration. These figures are very consistent with the age, tenure, and occupational profile of the workers in our sample.
taking a risky job. However, when we look at the wage regression results for between-job wage growth, we see that workers in higher risk (higher payoff dispersion) firms are taking lower starting salaries (or pay cuts) to move to more innovative firms (Table 6). Furthermore, if we hold constant the quality of workers, as measured by their previous pay, higher starting salaries in innovative firms reflects the hiring of talent, not a compensating differential for risk. Indeed, software workers may be seeking risk; they want to gamble some of their human capital wealth by taking risks at firms that might payoff big later. High talent people seek jobs in firms where they will be compensated highly if they succeed. The key lies in the nature of the risk taking. People may request high base pay for risk when they work in markets that have a lot of idiosyncratic risk, but highly talented individuals may increase their preference for incentive pay and risk-taking when the outcome arises not from noise, but from talent and effort.

Third, we conclude above that “loyalty” pays more than job-hopping. This is a very clear conclusion in our data: wage growth occurs after several years within a firm. There is still job hopping in software, but it pays in the long run, not in the short run. That is, one would not expect high-frequency job hopping (as in job changes every year). In order to earn returns from changing jobs, the workers must stay with the firm and work on the innovations, often in a team effort. Going beyond software, we would conclude that in markets where skills matter, workers must stay with the firms long enough to work with others and produce innovations.

---

37 In all our regressions, we find that pay is increasing in the amount of worker “churning” or gross turnover at firms. The inclusion of worker churn as a control helps to capture effects that may be associated with compensating differentials for risk taking, since turnover can be thought as a proxy for job security. That the main results are robust to the inclusion of this control provides evidence in support of the finding that the greater product payoff dispersion associated with higher earnings reflects firms’ efforts to attract and retain highly talented workers rather than to compensate for risk.

38 Using data on one large software firm, Russell (2005) finds that within the firm, base pay levels, bonuses, and options are highly correlated across individuals, reflecting the fact that more able workers have higher pay of every kind than less skilled workers. Under some circumstances, we might expect there to be less incentive pay in firms operating in high variance product markets. Indeed, in a tournaments model of incentive pay, increasing the amount of noise or luck reduces the use of incentive pay (Lazear and Rosen 1981). In our model, variance in payoffs could arise in part from idiosyncratic shocks representing noise or luck, but it also arises because some firms hire more talented people who select more successful products and should have pay tied to performance. In the data, we cannot identify whether the variance in the payoff arises from luck or effort, but our model of innovation proposes that it is high skill that produces high payoffs, so the coefficient \( \sigma^p \) should be positive as opposed to negative. Prendergast (2000, 2002) also points out that higher risk environments may have more performance-based pay because the cost of determining what inputs to monitor in such environments is greater. Since we cannot identify the source of the variance in payoffs and we do not have time-series data on product-specific variances or firm-specific variances, we turn to the data to determine the sign. For related empirical models of risk-pay incentive relationships, see Baker and Hall (2004), Core et al. (2003), Ittner et al. (2003), Murphy (1986), Schaefer (1998), and Wulf (2005). For excellent reviews of the literature on the subject, see Hallock and Murphy (1999) and Murphy (1999) for CEO pay.
Fourth, in some respects, our model is a one-dimensional model of talent. Since the output is innovation, the talent is the ability to innovate in project creation and selection. This talent is likely to reflect multiple skills—programming, management, etc.—all of which produce high compensation in return for the innovation. Thus, we have a simple model of one skill and the sorting to firms that value it. We have therefore avoided a matching model in which people have different comparative advantages and there can be stars sorting to all firm types. The reason we model one skill type is that we have one objective—to innovate—and innovating happens to produce the highest value added and highest wages.\textsuperscript{39} We are relying on the empirical evidence in our Software Census Survey data to show that, on average, ‘explorer’ firms that are doing the innovating (as in gaming) have higher value-added per worker than ‘exploiter’ firms that remain successful at exploiting previous innovations (as in mainframe software).\textsuperscript{40}

7 Conclusion

The process of innovation in the U.S. economy is fundamentally dependent on firms employing and rewarding highly talented workers. This paper draws a link between the product market strategies of firms and the structure of compensation. We argue innovation is a process in which workers create or select new projects. Therefore, firms that operate in product markets that have high potential returns to innovation—or a high rightward skew in the innovation gains in their product class—should select talent carefully and pay workers highly for these skills. To test this link between product market innovation and skill demand, we assemble panel data on individuals as they move across firms and link it to their firm’s product market strategy using the Census of Software firms. We show that software firms that operate in product markets with highly skewed returns to innovation, or high upside gains to innovation, are more likely to attract and pay for highly talented workers. Such firms do so first by paying more up-front in starting salaries to attract skilled employees, and second by rewarding talented workers handsomely for loyalty. These striking effects are robust to the inclusion of a wide range of controls for both

\textsuperscript{39} If there were two skills, such as skill at innovating and skill at customer-driven design, and workers sort to their optimal sector based on comparative advantage, the workers who are skilled at innovating may well earn less in the sectors that value customer-driven design. Such a model would then show that we are underestimating the returns to innovating and underestimating the returns to customer-driven design.

\textsuperscript{40} For more explanation of the ‘explorer’ innovative versus ‘exploiter’ concepts, see the strategy literature (Saloner, Shepard, Podolny, 1999). Of course, there certainly can be some exploiter firms that have value added that is higher than many explorer firms: our data shows that exploring has higher payoff on average than exploiting.
worker and firm characteristics, including variables capturing rent-sharing when a firm is currently successful as well as proxies for other types of risk.

Though we focus on the software industry, our model and findings should generalize to other industries in which firms employ knowledge workers and face uncertainty in the probability of success on any given project. Our results documenting a link between income variance and innovation also complement the literature on income inequality, changing skill demand, and economic growth. Recent research suggests that returns to skill have been increasing within as well as across occupations and industries, and furthermore that increases in earnings inequality in recent decades have been driven largely by changes in the upper as opposed to lower tail of the income distribution (Autor et al. 2003; Autor et al. 2005, 2006; Lemieux 2006). In addition, Lemieux et al. (2007) show that rewarding high skilled workers with performance pay accounts for much of the increase in income inequality at the top end of the pay distribution. Our results for the software industry speak to these broader patterns. We introduce the product market characteristics of software firms as an index of skill demand: highly innovative high-technology firms pay a premium for talent, contributing to a highly skewed distribution of earnings. We cast this inequality in a positive light, showing that high variance in earnings goes hand-in-hand with innovative activity in dynamic and risky markets. To the extent that these markets have been and will continue to be a source of growth in the economy, our research makes important contributions to our understanding of not only firm human resource practices and product market strategies, but also patterns of income inequality and economic development.
References


Appendix

Figure A1:
Difference in Predicted Starting Earnings at Observed Maximum and Minimum Product Market Risk Levels
Experienced Workers 21–44 Earning $50,000+, Other RHS Variables at Means

Figure A2:
Difference in Predicted Experienced Earnings at Observed Maximum and Minimum Product Market Risk Levels
Experienced Workers 21–44 Earning $50,000+, Other RHS Variables at Means

RHS variables include product payoff dispersion, log revenue per worker, employment, employment squared, firm age dummies, firm employment growth rate, location dummies, and time dummies. Variables set at means include log revenue per worker, employment, employment squared, and firm employment growth rate. Assumes firm age of 6–10 years; that the firm is located in a high density, high education, and industrially diverse county; and that the worker accessed in the second quarter.
Figure A3:
Difference in Predicted Minimum Ending Earnings at Observed Maximum and Minimum Product Market Risk Levels

Experienced Workers 21–44 Earning $50,000+, Other RHS Variables at Means

RHS variables include product payoff dispersion, log revenue per worker, employment, employment squared, firm age dummies, firm employment growth rate, location dummies, and time dummies. Variables set at means include log revenue per worker, employment, employment squared, and firm employment growth rate. Assumes firm age of 6–10 years; that the firm is located in a high density, high education, and industrially diverse county; and that the worker separated in the second quarter.

Figure A4: Coefficients on the Product Payoff Dispersion across the Earnings Distribution
For Experienced Earnings

Worker controls include quadratics of tenure at job, tenure in industry, and age, fully interacted with each other and with appropriate left and right censoring dummies. Firm controls include a quadratic in (log) firm employment, dummies for firm age (<6 years, 6-10, 11+ years), the net employment growth rate, and a dummy for whether the firm is in a high density/high education/industrially diverse county. Controls also include time dummies for quarter of separation and/or quarter of accession as appropriate. Based on LEHD data for ten states.