I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12320

Incentivizing Learning-By-Doing: The Role of Compensation Schemes

Joshua Graff Zivin
Lisa B. Kahn
Matthew Neidell

APRIL 2019

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12320

Incentivizing Learning-By-Doing: The Role of Compensation Schemes

Joshua Graff Zivin
UC San Diego and NBER

Lisa B. Kahn
University of Rochester, NBER and IZA

Matthew Neidell
Columbia University, IZA and NBER

APRIL 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Incentivizing Learning-By-Doing: The Role of Compensation Schemes[1]

In this paper, we examine the impact of pay-for-performance incentives on learning-by-doing. We exploit personnel data on fruit pickers paid under two distinct compensation contracts: a standard piece rate plan and a piece rate plan with an extra one-time bonus tied to output. Under the bonus contract, we observe bunching of performance just above the bonus threshold, suggesting workers distort their behavior in response to the discrete bonus. Such bunching behavior increases as workers gain experience. At the same time, the bonus contract induces considerable learning-by-doing for workers throughout the productivity distribution, and these improvements significantly outweigh the losses to the firm from the distortionary bunching. In contrast, under the standard piece rate contract, we find minimal evidence of bunching and only small performance improvements at the bottom of the productivity distribution. Our results suggest that contract design can help foster learning on the job. This underscores the importance of dynamic considerations in principal-agent models.

JEL Classification: J33, J43
Keywords: contracts, learning-by-doing

Corresponding author:
Matthew Neidell
Department of Health Policy and Management
Mailman School of Public Health
Columbia University
722 W. 168th St.
New York NY 10032
USA
E-mail: mn2191@columbia.edu

[1] We thank Arthur Campbell, Richard Holden and seminar participants at Colgate, Cornell University, University of Arizona, University of California, Los Angeles, University of Chicago, University of New South Wales, University of Miami, and Vanderbilt.

1. Introduction

The importance of performance-based pay in employment contracts when worker effort is imperfectly observable has been well understood for more than three decades (e.g. Harris and Raviv, 1979; Holmstrom, 1979; Grossman and Hart, 1983). However, with the notable exception of Gibbons and Murphy (1992), surprisingly little attention has been paid to dynamic concerns in these principal-agent models. At the same time, the learning-by-doing model first proposed by Arrow (1962) and central to many modern economic growth theories (Romer, 1994) has largely neglected the role that firm-level incentives can play in fostering learning by employees. In this paper, we bring these two distinct strands of literature together to show these gaps are empirically consequential.

In particular, we exploit a unique panel dataset on individual worker productivity to examine the role of performance-based pay on job-based learning. Our data comes from a grape and blueberry farm that utilizes two distinct contracts to compensate workers for their efforts. The contract for grape pickers pays a fixed wage up to a productivity target and a piece rate thereafter. The contract for blueberry pickers pays a fixed wage up to a productivity target, a bonus for reaching the target, and a piece rate thereafter. We primarily focus on whether the stronger incentives embedded in the second contract translate into more or faster learning on the job. As such, we follow in the rich empirical tradition of studying fruit pickers to examine employee incentives and various aspects of agency theory (see Bandiera et al. 2011 for a review).

We first develop a conceptual model where workers take costly hidden effort to produce output. Workers vary in their convex effort cost such that some will find it optimal to produce at a low effort-low reward point, while others produce at a high effort-high reward point. This model provides two empirical insights. First, when there is a discrete bonus, as in the case for the blueberry contract in our empirical setting, some workers will find it optimal to bunch just to the right of the bonus point. Second, while all workers have an incentive to lower their effort costs, i.e., learn-by-doing, the incentive is greater under the blueberry contract, which offers the bonus payment if a worker crosses the threshold.

Consistent with our conceptual model and the literatures on salespeople and executive compensation (Healy, 1985; Oyer, 1998; Larkin 2014, Freeman et al., 2019), we find that the contract that includes a bonus leads to strategic heaping around the threshold. Importantly, this behavior changes as workers gain more experience. More seasoned workers are better able to reach the threshold and engage in heaping. We observe no such heaping or changes in heaping for workers under the commission only contract (grape pickers). For example, we find that blueberry pickers increase their likelihood of hitting the piece rate threshold by 15 percentage points (80%) over their first 10 days on the farm, while grape pickers see very little improvement. This heaping is potentially costly to the firm since small changes in output are accompanied by substantial changes in payout to workers.

Further, we find that the productivity gains are not limited to only those workers in striking distance of the bonus payment. In fact, we see improvements with tenure throughout the output distribution. In contrast, in the standard piece rate contract, we only see learning at the bottom of the productivity distribution, consistent with the notion that fears of termination drive learning for these workers. Furthermore, the bonus payment induces productivity gains that are sufficient to more than offset the costs induced by inefficient heaping. A simple back of the envelope calculation based on empirical estimates of learning-by-doing under the blueberry contract suggests that the learning induced from day 3 to day 10 of tenure generates an additional $60 in net revenue per worker, or approximately $8.50 per day.

Thus, the large bonus payment appears to be instrumental in fostering learning-by-doing, making this seemingly flawed contract a clear winner from the firm's perspective.[2] That said, it is important to note that crop and contract type are collinear in our setting, opening the possibility of alternative hypotheses centered on difference across crops rather than contracts. One obvious candidate is that there is simply more scope for learning in the harvest of blueberries than in grapes. While we have not found any evidence in the agronomic or related literatures that might support this contention, we cannot completely rule it out. Nonetheless, we are reassured by the large degree of heterogeneity in base-level performance, and the fact that workers at the bottom end of the performance distribution of grapes improve with experience, both of which highlight that the potential scope for learning does not appear, prima facie, limited by crop type. Another possibility is that better workers are drawn to blueberries. We are able to dispel this concern more directly -- we find that within-crop learning behavior is very similar for workers observed picking in both crops.

By showing that performance-based pay has important implications for learning-by-doing, this paper contributes to two fundamental, but heretofore, disconnected economic literatures. Our results imply that contracts that appear sub-optimal in the short run may be efficient in the long run. This suggests that contract design has an important role to play in fostering learning on the job, a topic that, outside of direct R&D investment, has largely been ignored in the learning-by-doing literature (See Thompson, 2012 for a review). It also suggests that, in settings where firm- or industry-specific human capital exist (Becker, 1962), contract design can overcome not only moral hazard in effort but also moral hazard in learning.

The remainder of the paper is organized as follows. Section 2 presents a simple conceptual model of worker output under two distinct performance-based pay contracts that resemble those that form the basis of our empirical work. Section 3 describes our study setting and data.

[2] In the spirit of the Holmstrom-Milgrom multitasking model (1994), we view heaping and learning as complements in production.

Section 4 describes our econometric strategy and results. Section 5 examines mechanisms and robustness. Section 6 offers some concluding remarks.

2. Conceptual Model

In this section, we develop a simple conceptual model that mirrors some of the key features of our empirical setting. Workers face one of two types of compensation schemes. Under both contracts, workers are paid a fixed base wage plus additional pay based on performance after reaching a productivity target. The first contract, the "commission contract," pays workers a piece rate for each unit of output that exceeds the threshold. In our empirical setting, this contract corresponds to a piece-rate subject to a minimum wage; it is also representative of a wide array of compensation schemes that include a sales commission or profit-sharing arrangement with some form of minimum salary guarantee. The second contract, the "commission plus bonus contract," is similar in structure but also includes a one-time bonus when workers reach their productivity target. Such bonuses are commonplace in salesperson and executive compensation schemes where they are justified as a tool to help workers with goal setting (Joseph and Kalwani, 1998; Healy, 1985).

2.A. Commission Contract

We begin with an examination of equilibrium output levels under the commission contract. Equation (1) describes total take-home pay as a function of the minimum guaranteed salary, w, output, q, the pay-for-performance threshold, ϕ, and the piece rate, p. Here, workers receive a piece rate on all output above the productivity target, ϕ. We assume that pay also varies with output below the threshold ϕ at a rate $f(q)$. This allows for the possibility that additional output increases the probability that a worker can keep his job (Lazear 2000).[3] Let $f'(q)$ represent the implicit wage rate per piece below the productivity target. The take home revenue for a worker is the sum of three terms: (1) the minimum salary guarantee plus (2) some benefit from a reduction in the likelihood of termination that depends on q plus (3) the piece rate for any units that exceed the threshold.

[3] We assume for simplicity that the termination threat is zero for those who exceed the productivity target.

$$Take\ Home\ Revenue = w + \int_0^{\phi} f'(q)dq + \int_{\phi}^{\infty} p\, dq \qquad (1)$$

Let $c_j(q)$ denote the costs of producing q units of output for a worker of type *j*, where costs are convex in output levels, such that $c_j'(q) > 0$ and $c_j''(q) > 0$ for all *j.* Figure 1 illustrates examples of equilibrium outcomes for two different cost functions, a high and low cost of effort. MR denotes the workers' marginal wage (revenue) schedule, which is upward sloping until the productivity target, and equal to *p* thereafter*,* reflecting the constant commission.[4]

High effort costs: We define workers with high effort costs as those whose marginal cost of effort (MC_H) increases steeply, thereby crossing the *MR* curve only once at a point below ϕ. An individual of this type receives the minimum salary guarantee (*w*), no commission, and faces a non-zero probability of termination. This equilibrium is illustrated in Figure 1, where *x* indicates the productivity level such that $f'(x) = c_H{}'(x)$ for a worker of high effort cost type.

Low effort costs: We define workers with low effort costs as those whose marginal cost of effort curve (MC_L) increases slowly with output levels such that it intersects the marginal revenue curve only once and strictly above the productivity target ϕ. An individual of this type will be paid the wage guarantee as well as additional compensation for each unit of output that exceeds ϕ. This equilibrium is illustrated in Figure 1, where *y* indicates the output level for a worker of this type, and $p = c_L{}'(y)$.

2.B. Commission Plus Bonus Contract

We extend our analysis to an examination of equilibrium daily output levels under the commission plus bonus contract. Revenue under this contract is nearly identical to the commission only case, with one important exception. Workers receive a one-time bonus

[4] The figure draws one example of f'(.) where f''(.) is a positive constant below the piece rate threshold. In reality, f(.) could be increasing, decreasing or non-monotonic in output. The general intuition that follows is unaffected by this assumption, although the relative importance of effort cost type will vary across specifications.

payment b when they hit the threshold ϕ. As such, total revenue under the commission plus bonus scheme is characterized by the following piecewise function ($\mathbf{1}$ is an indicator function equaling one if output exceeds ϕ):

$$Take\ Home\ Revenue = w + \int_0^{\phi} f'(q)dq + \int_{\phi}^{\infty} p\, dq + \mathbf{1}\{q \geq \phi\} * b \qquad (2)$$

Let $c_j(q)$ denote the costs of producing output level q for a worker of type *j*, as before. We describe three types of equilibrium outcomes, each illustrated in Figure 2.

The high and low effort costs equilibria are similar to those under the commission only case. The former corresponds to output levels in which workers are paid the minimum wage guarantee and face a non-zero probability of termination. The latter corresponds to output levels in which workers are paid the guarantee as well as additional compensation for each unit of output that exceeds ϕ. These are illustrated as point x and y in Figure 2, respectively.

Moderate effort costs: The more interesting case under this contract occurs when workers have a marginal cost structure between the high and low effort costs. We define these "moderate effort costs" workers as those whose marginal cost of effort curve (MC_M) crosses the marginal revenue curve at two distinct points, one below and one at the target productivity threshold ϕ. Points z and z' in Figure 2 illustrate these output levels. In this case, we need to evaluate which outcome is optimal.

As indicated in Figure 2, let *k* denote the marginal cost (in dollar terms) of producing output z' (which equals the target bonus output level ϕ). A worker will choose to produce at this level if and only if the following holds:[5]

[5] For simplicity and without loss of generality, we describe the net marginal cost curve of moving from z to the bonus payment as an integral from z to ϕ. Since we model the bonus payment as one that is paid precisely at output level ϕ, the upper integrand on the left-hand side of (3) should technically be ϕ-ε, where epsilon is an arbitrarily small unit of output.

$$\int_z^\phi [c'(q) - f'(q)]\, dq \leq [\mathrm{b} - k] \qquad (3)$$

That is, a worker needs to produce $(\phi - z)$ additional units of output to hit the threshold. A worker will only choose this higher output level when the costs incurred from those extra units of output are sufficiently compensated by the lump-sum bonus payment, b, plus any additional benefits in job stability (f'(.)).[6]

We illustrate the heaping phenomenon in Figure 3. Let q^* correspond to the level of output that defines equation (3) with equality for the marginal cost curve, MC^*, which intersects the MR curve at q^*. Moderate effort cost workers with a higher marginal cost curve than MC^* (i.e. those whose marginal cost curve intersects MR below q^*), such as MC_j, produce less than ϕ output and receive the minimum salary guarantee. In contrast, moderate effort costs workers with a lower marginal cost curve than MC^* (i.e. those whose marginal cost curve intersects MR above q^*), such as MC_k, will heap at the target productivity level ϕ. Equilibrium output across all worker types is highlighted in bold, with a heaping zone defined by the area between q^* and ϕ. Any individual within that zone earns a bonus for reaching the threshold that exceeds the excess costs of reaching that point, thus leading those who would have otherwise produced output between q^* and ϕ to heap at the target productivity threshold. While conceptually similar to the heaping found in the tax notch literature (e.g. Saez, 2010; Kleven et al., 2011) and that found for sales workers in Oyer (1998), heaping in our context reflects a genuine increase in output for those near the productivity target rather than misreporting or a simple time shifting of output.

2.C. Learning by Doing

[6] For simplicity, we limit our attention to the case where the marginal cost curve intersects the marginal revenue curve twice. Given our functional form assumptions, it is possible (if not probable) that the marginal cost curve could intersect the marginal revenue curve three times. In this case, the logic outlined below on heaping is altered. Workers will either produce output below the target threshold or above it (but not heaped), depending on which generates the larger amount of total net revenue (as per above).

One of the motivations for this paper is to examine learning by doing. In particular, we are interested in the case where workers learn about how to improve their execution of a task through experience.[7] For simplicity and without loss of generality, we simply treat learning by doing in our model as shifting $c_j(q)$ downward for any given worker as a function of time spent on the job.[8] Thus, some high effort cost workers may become moderate and eventually low effort cost workers over time.

Both contracts in our setting should induce learning: low cost of effort workers are better off than high cost workers because workers extract at least part of the rent associated with their output through increased piece rate payments above the threshold or reductions in termination probabilities below it.[9]

Importantly, however, these learning incentives are stronger under the bonus contract. This stems from the fact that, as we have shown, workers in a certain output range find it optimal to heap (illustrated in Figure 3). These workers are better off heaping under the commission plus bonus contract than not heaping under the commission only contract, thanks to the large bonus they receive under the former contract. As such, a worker initially below output level q^* has an added incentive to improve in the commission plus bonus contract, relative to a worker in the commission only contract whose improvements only impact their firing probability. The larger the bonus in the former contract, the stronger the incentive to lower effort costs so that they can reach the threshold that triggers its payment.

[7] Note that this is distinct from another branch of the learning-by-doing literature that is more focused on workers or firms learning about their time-invariant ability to carry out a task. This includes the match-quality models first introduced in Jovanovic (1979) and the employer learning models explored in Farber and Gibbons (1996), Altonji and Pierret (2001), Lange (2007), and Kahn and Lange (2014).

[8] The underlying mechanisms that drive shifts in c(q) is deliberately unspecified. The workhorse models of productivity evolution often have a human capital framework in mind (Becker 1964, Mincer 1958, and Ben-Porath 1967), although others employ a Bayesian framework where feedback to workers on the returns to their effort/technique is noisy and learned through repeated draws from a known distribution (Jovanovic and Nyarko, 1995). Our conceptual model is sufficiently flexible to accommodate a wide range of mechanisms for learning.

[9] Low cost of effort workers exert more effort but are rewarded more for their trouble. The envelope theorem tells us that this combination is preferable to the low-output/low-rewards point that a high cost of effort worker chooses.

We should then expect to see increases in heaping behavior as workers gain experience under the commission plus bonus contract. Even workers who begin with high effort costs and an equilibrium output far below the heaping threshold still have an incentive to learn. Within the context of the model, these workers steadily increase the probability that they eventually earn the bonus as they accumulate skills through their increased experience on the job. Outside the context of the model, the salient bonus may serve as a motivator in and of itself. [10] Either way, the bonus payment should accelerate learning all along the productivity distribution, relative to the commission contract.

Thus, our basic model predicts two empirical regularities that can be summarized as follows:

1. The commission plus bonus contract leads to heaping (as illustrated in Figure 3).[11]
2. If most workers begin as high effort cost types, we should see faster improvements in output as workers gain experience under the commission plus bonus contract. These improvements should result in an increase in the rate of heaping with job tenure, as well as a steeper tenure-productivity gradient below the heaping point.[12]

Importantly, the increased learning-by-doing associated with the commission plus bonus contract may help explain the very existence of this seemingly suboptimal type of contract. After all, the heaping that occurs due to the threshold bonus is costly to firms since the payment to workers changes non-trivially around the threshold for marginal changes in productivity. To a first-order approximation, the costs of this heaping can be expressed as the marginal costs of wages to those heaped at the threshold. Firms should only be willing to incur those costs if the features of this contract generate enough productivity gains elsewhere in the

[10] Psychologists have argued that salient bonuses can be important for directing attention, pulling task-relevant knowledge into awareness, and increasing persistence (Locke et al., 1981; Locke and Latham, 1990). This may also generate a small amount of heaping in grapes.

[11] Although we do not model this explicitly, if focal points provide psychological incentives to reach the piece rate threshold, then we may see heaping in grapes. Since this psychological effect should apply to both grapes and blueberries, our prediction would then be that we expect to see more heaping in blueberries relative to grapes since the former also has a strong financial incentive for reaching that threshold.

[12] More generally, the amount of heaping that results from additional experience will depend on the initial distribution of worker types and the specific form of the learning process.

distribution. This is particularly relevant for those below the productivity target where each unit increase in output is pure rent for the firm because workers receive a guaranteed fixed wage not directly tied to output.

3. Data and Setting

The dataset used in this paper is personnel data of harvest workers from a farm in the Central Valley of California.[13] To protect the identity of the farm, we can only reveal limited information about their operations. The farm, with a total size of roughly 500 acres, produces blueberries and grapes during the warmer months of the year. Our dataset covers the growing seasons of 2009 and 2010. In general, blueberries are picked in May and June, while grapes are picked in August and September. As detailed below, all workers face some type of performance-based pay arrangement.

Our data consist of a longitudinal file that follows workers over time by assigning unique identifiers based on the barcode of employee badges. It includes daily information on the total number of pieces harvested by each worker, the location of the field, the type of crop, the terms of the piece rate contract, time in and out, and the gender of the worker.[14] Our final dataset follows an unbalanced panel of roughly 1,300 workers over 101 days of harvest activities.[15] Data quality is extremely high, as its primary purpose is to determine worker wages.

[13] This dataset was first used by Graff Zivin and Neidell (2012) to study the impact of air pollution on worker productivity.

[14] We obtain these data from a unique arrangement with an international software provider, Orange Enterprises (OE). OE customizes paperless payroll collection for clients, called the Payroll Employee Tracking (PET) Tiger software system. It tracks the progress of employees by collecting real-time data on attendance and harvest levels of individual farm workers to facilitate employee and payroll management. The PET Tiger software is installed on handheld computers used by field supervisors. At the beginning of the day, supervisors enter the date, starting time, and the crop being harvested. Each employee clocks in by scanning the unique barcode on his or her badge. Each time the employee brings a piece, his or her badge is swiped, recording the unit and time. Data collected in the field is then synchronized to a host computer, which facilitates the calculation of worker wages.

[15] To eliminate potential concerns about the interaction of contract type and free-riding, we limit our attention to the crops in our dataset that are paid based on individual (rather than team) output.

Worker output is measured in "pieces," where blueberries are collected in buckets and grapes in boxes. One piece is roughly equivalent to 5-pounds of fruit, which corresponds to approximately 400 grapes or 1000 blueberries. At the time of our study, the wholesale prices were, on average, \$2.83 per pound for blueberries and \$0.33 per pound for grapes.[16] In general, our productivity measures are available at the daily level, although we have hourly productivity measures for a small subset of our sample.[17]

The farm offers two distinct performance-based compensation contracts depending on the crop being harvested for reasons that are largely a historical artefact.[18] Grape harvesters face a commission scheme, as described in our conceptual model. Workers earn \$8 per hour (the minimum wage in California during the time of our study) until their daily output reaches the piece rate threshold of 32 pieces. At that point, the harvester earns an additional bonus per box picked. The piece rate was \$0.30 per box above the threshold in 2009 and \$0.35 in 2010.

For blueberries, the compensation plan resembles the commission plus bonus scheme described earlier. Workers continue to earn \$8 per each hour worked until their daily output reaches the piece rate threshold of 25 pieces. Upon reaching the threshold a worker is paid a sizable one-time bonus equal to \$12. She then earns a piece rate of \$0.50 for each additional piece harvested beyond the threshold. More details on the dataset and piece rate schedules can be found in Graff Zivin and Neidell (2012).

Besides the one-time bonus, the contracts differ in the threshold at which the piece rate kicks in and the size of the piece rate. Though the primary task performed by grape and blueberry

---

[16] These figures are the average of grower prices for the 2009 and 2010 growing seasons. For blueberries, see table 12 of http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1765. For grapes see Table A-12 of 2011 Yearbook at http://usda.mannlib.cornell.edu/MannUsda/viewStaticPage.do?url=http://usda.mannlib.cornell.edu/usda/ers/89022/2011/index.html.

[17] We obtained this hourly sample for a fraction of workers on the first blueberry harvest as the farm experimented with the new payroll collection software.

[18] In personal communications with the farm, they have indicated that they have offered these two distinct contracts for as long as anyone could remember and that they have maintained them to ensure consistency.

pickers is similar, grape boxes are easier to fill than blueberry buckets given their size. This fact is reflected in their higher threshold, lower piece rate, and lower price per pound.

We restrict our sample to those workers who eventually complete at least 10 days with the farm. We focus our attention on this subsample for the analysis so that worker composition does not influence our learning-by-doing results. The online appendix presents all of our main results using all workers, regardless of tenure. All of our conclusions hold, both qualitatively and quantitatively, with or without this restriction.

Table 1 provides some basic summary statistics. The average grape worker harvests 33 boxes per day, with 55% of worker-days reaching the target productivity level at which performance-based pay begins. Consistent with the additional fruit required to fill a bucket (and the corresponding pay schedule), blueberry harvesters produce fewer pieces and are less likely to reach their production target. Average output is 20 pieces per day, with the blueberry productivity target reached on 36% of worker-days. Despite these differences in output levels, realized wages are fairly similar across crops, with blueberry pickers averaging about $3.80 more per day than grape harvesters. The variance in earnings is higher for blueberry pickers than for grape pickers, commensurate with the stronger incentive component for the former.

Since much of our analysis focuses on learning-by doing, Table 1 also provides some basic statistics on job tenure within our dataset. The average tenure of a grape (blueberry) worker at a point in time is 14 (12), and the average completed tenure is 28 (24) days at the farm, although these range from 10 to 70 days.[19] Finally, it is worth noting that approximately 7% of workers have experience in both grapes and blueberries, a sample that we will exploit to understand the role of worker selection into each compensation contract. [20]

[19] While not all workers start on the first day of the growing season, 57% of grape pickers and 66% of blueberry pickers are employed within the first 10 days of the given harvest.

[20] For our main analyses, we restrict our attention to the time spent in a worker's first crop since it is unclear how much learning ports across crops (a point we explore separately). This restriction leads us to drop 3 percent of our blueberry worker day observations and 15 percent of our grape worker day observations.

4. Econometric Strategy and Results

Inspired by our conceptual model, we start by comparing output across the two contract types to test whether workers respond to the sharp nonlinear payout in the commission plus bonus contract by heaping just beyond the threshold. Next, we explore learning-by-doing by examining the evolution of worker productivity as a function of experience. We initially limit our focus to changes in the likelihood that workers reach their productivity targets, but also explore learning-by-doing throughout the productivity distribution. As noted above, we perform our analysis on the sample of workers who survive at least 10 days with the farm to limit concerns about changing sample composition. All tables and figures are replicated in the online appendix for the unrestricted sample.

4.A. Heaping

Recall that the commission plus bonus structure creates strong incentives for heaping just beyond that threshold. Figure 4 plots the histogram of daily worker output separately by crop, where the vertical line indicates the piece-rate threshold. Focusing first on blueberries (left), the heaping around the bonus payment threshold (25 pieces) is immediately evident. To the left of the threshold, the mass is generally normally distributed, discontinuously jumps up at 25 pieces, and steadily declines thereafter. In contrast, the distribution of productivity for grape harvesters (right) is relatively smooth throughout.

This visual analysis is supplemented by a more formal test based on a recently developed non-parametric test for heaping for discrete data (Frandsen, 2017). This procedure was initially designed to test for manipulation in the running variable in a regression discontinuity design, but can be readily applied to testing for heaping, as in our context.[21] The test is based on a smoothness approximation around the threshold, but only relies on mass at the threshold and immediately adjacent. Consistent with the theory and the visual evidence in the figures, heaping just past the target productivity threshold is statistically significant and large in

[21] This test is the discrete analog to that of McCrary (2008), which breaks down because the local linear regression requires a continuous support near the threshold. Given the discrete nature of worker output, especially around the thresholds, we opt for the test by Frandsen.

magnitude in blueberries. Heaping for grapes is also statistically significant (p=0.015) – perhaps due to workers responding to the salience created by the threshold (Locke et al., 1981, Locke and Latham 1990) – but the magnitude of the displacement at the threshold is markedly smaller than the one found in blueberries. The number of blueberry workers just beyond the threshold increases by roughly 9 percentage points, whereas the corresponding measure for grape workers is 1 percentage point.

4.B. Learning to Heap

We next explore learning-by-doing by examining how heaping evolves with tenure. We equate learning-by-doing with days worked rather than pieces produced. Both measures are potentially valid constructs: while workers learn from previous output, they also learn how to pace themselves throughout the day, a concept captured by days worked. We opt for days worked (or tenure) as the work experience variable because the particular day of work (conditional on the number of completed days) is exogenous with respect to the contract while the number of pieces is not.[22]

Under the blueberry contract, we expect workers to become more effective at reaching the threshold as they gain experience on the job relative to grapes, where there is no salient bonus. We explore experience profiles using figures similar to above but plotting them separately by tenure with the farm in 2-day groupings.[23] In Figure 5, which plots blueberry pickers, we see strong evidence of increased heaping with tenure. Few blueberry pickers cross the threshold to reach the bonus on their first two days of work. By days 3-4, the amount of heaping has clearly increased, and continues to do so over the next several days. By the end of their second week of work (roughly day 10), the degree of heaping appears relatively stable.

[22] We opt for estimating simple productivity growth models that are linear or quadratic in days worked, given our setting and our key questions of interest. Other empirical work on human capital accumulation estimates more complicated structures, including random productivity shocks and idiosyncratic persistent growth rates (see for example, Abowd and Card 1989, Baker 1997, Guvenen 2007, Hause 1980, and McCurdy 1982, among many others).

[23] We opted for 2-day groupings over 1-day for ease of exposition, though results are quite similar when disaggregated.

Figure 6 shows the same plots for grape pickers. Consistent with Figure 4, we see little evidence of heaping regardless of tenure with the farm. These two figures are consistent with our theory that heaping is a learned behavior that is largely limited to the blueberry contract, which offers the bonus payment upon reaching the productivity target.

To test for the magnitude of these differences, we estimate tenure profiles for the probability that a worker crosses the threshold on a given day, and allow these profiles to vary by crop. We include controls as indicated below, experiment with different functional forms for the tenure profile, and cluster standard errors by worker. Regression results are summarized in Table 2.

The first column includes just three regressors: a linear tenure profile, an indicator for crop, and their interaction. We find that while both blueberry and grape pickers improve with additional days of tenure, blueberry pickers improve faster. Grape pickers increase their probability of meeting the piece rate threshold by 0.4 percentage points with each day of tenure, off their base probability of 49% on their first day (the constant). The coefficient on the blueberry indicator reveals that blueberry pickers are, on average, 31 percentage points (63%) less likely to meet the piece rate threshold, compared to grape pickers. However, with each day of tenure, blueberry pickers increase their probability of passing the threshold by an additional percentage point beyond grape pickers, which is a roughly 3.5 times greater rate. This difference is statistically significant at the 1% level. Our estimates imply that after 10 days, blueberry pickers increase the likelihood of reaching the target by 15 percentage points (80% off their base of 18%), while the likelihood for grape pickers only increases by 4 percentage points (8% off their base of 49%).

The remaining columns of table 2 show that results are largely unchanged when we control for tenure and its interaction with the crop more flexibly or when we control for worker gender and characteristics of the day of observation.[24] Furthermore, the differential tenure profiles are always statistically significant at the 1% level. With the quadratic-in-tenure specification and full controls (column 4), we find that blueberry pickers increase their likelihood of meeting the piece rate threshold by nearly 14 percentage points (116%) over their first 10 days, while grape pickers improve by only 6 percentage points (12%).

4.C. Does Total Output Increase with Experience?

The increased heaping with tenure induced by the blueberry contract indicates one of two phenomena: pickers learn how to produce more efficiently or pickers learn how to better game the contract. The former is clearly more beneficial to the firm, while the latter could be particularly costly. To better understand the relative role of each, we expand our focus beyond the threshold to examine productivity throughout the distribution.

We begin by noting that a simple visual inspection of the histograms in Figure 5 suggest that the increases in productivity amongst the more seasoned workers is not simply about gaming since the entire productivity distributions appear to shift rightward as workers gain experience. These shifts are especially pronounced in the first two rows, as workers make visible improvements in productivity.

To further probe learning under the blueberry contract, we turn our attention to the starting performance of workers who eventually hit the threshold. If most of the workers who eventually hit the threshold had a starting performance just below the threshold, then it would suggest that workers are largely learning to game the contract and all productivity gains will be quite marginal. If, on the other hand, many of those that eventually reach the threshold had a starting performance far from the threshold, then we would infer that workers have made substantial improvements in their productivity.

[24] Gender controls include indicators for male and female, with missing gender as the omitted category. Date controls are day of week, week of harvest, and first and last day of harvest fixed effects. Including these controls does flatten the tenure profile for blueberries somewhat, which reflects the fact that tenure is correlated with the period in the harvest when crops are more plentiful (Stevens, 2017).

Figure 7 illustrates the relationship between initial productivity and the likelihood of reaching the bonus threshold. In this figure, we plot performance distributions for the first three days for blueberry pickers (left) and grape pickers (right). The histograms in the top panel show productivity for workers who do not reach the bonus threshold on days 10-12, and the bottom panel shows productivity for workers who reach or exceed the bonus threshold at least once on days 10-12.[25]

For blueberries, we find that workers who eventually reach the threshold come from all along the distribution, similar to those who do not reach the threshold on days 10-12. The distribution centers around 15 pieces (approximately 9 pieces below the threshold), indicating that these workers achieve substantial growth over their early days of experience. Further, blueberry pickers who eventually meet the threshold do not come from just below the threshold itself, as evidenced by the missing mass just to the left of the threshold. Turning to grapes, eventual threshold hitters are quite similar to the right of the threshold but there is less mass to the left of the threshold, suggesting the workers come from a higher part of the distribution to begin with. That blueberry workers experience considerable improvements in the lower tail of the distribution while grape workers do not provides further support for our contention that the learning for blueberry pickers is contract induced.

We can take a more systematic look at this learning by plotting the number of pieces picked at each decile of the output distribution by tenure separately for blueberries and grapes; we show this in figure 8, where darker lines represent workers with less experience. For blueberry workers (left), the lines monotonically shift upward with tenure, indicating that more experienced workers perform better than less experienced workers throughout the distribution. More specifically, a bottom performing experienced worker performs better than a bottom performing new worker, as does a middle-performing or top-performing experienced

[25] We choose day 1-3 and 10-12 windows to reduce noise in day-to-day performance. However, graphs restricted to single days or pairs of days look similar.

worker compared to middle or top performing new workers, respectively. In other words, each distribution stochastically dominates the tenure group just below it.

Note that most of the output distribution is below the piece rate threshold (shown by the horizontal line), for all experience levels. Thus, workers all along the distribution appear to improve, even those quite far from the heaping region around the threshold. Across the full tenure range plotted, growth in daily output for blueberry pickers ranges from 4 to 8 pieces, depending on the percentile. Since this figure is restricted to workers who survive at least 10 days, sample composition cannot explain the improvement of new workers.

For grapes (right) the pattern is quite different. Improvements over the first 10 days are small, between 1 and 2 pieces, depending on where they fall in the distribution. The largest improvements occur for the bottom two deciles. Grape pickers appear to experience less learning throughout the productivity distribution than their blueberry picker counterparts who face the commission plus bonus contract.

Table 3 provides further evidence on this pattern, but with a better understanding of statistical significance. We estimate quantile regressions of output at each quintile of the daily performance distribution. Output is normalized by crop to have mean zero and standard deviation one to account for the differences in the output distributions between grapes and blueberries. (We later explore other normalizations.) For each quintile, we regress output on a quadratic in tenure, a blueberry dummy, and interactions between the two. This allows us to assess whether improvements with tenure differ across crops at various points in the performance distribution. Regressions include full controls from table 2 and again cluster standard errors by worker.

The results indicate that output improves at the bottom of the performance distribution (20^{th} percentile) for both grape and blueberry pickers. Over the first 10 days, we estimate that the 20^{th} percentile improves by 0.25 of a standard deviation (about 2 pieces) for grape pickers. This improvement is statistically indistinguishable from that of the blueberry tenure profile; the bottom part of the performance distribution improves at roughly the same rate across crops.

At all other parts of the distribution, however, the performance-tenure profiles are steeper for blueberry pickers than for grape pickers. Further, these differences are statistically distinguishable at the 5% level or better. Over the first 10 days of tenure, blueberry pickers improve by roughly 0.3 to 0.4 standard deviations (or 2-3 pieces) at the 40th, 60th, and 80th percentiles. In contrast, the tenure profiles for grape pickers in this part of the distribution are essentially flat.

Figure 9 provides a graphical illustration of the results from Table 3 by plotting the fitted 10-day productivity changes based on the tenure and crop coefficients. We estimate profiles for each decile (rather than quintile) of the productivity distribution to better illustrate where the pattern emerges. We do this for 3 different outcomes: pieces (top left), standardized pieces, as in Table 3 (top right), and the "latent wage" of each crop (bottom left), which normalizes the outcome by the crop-specific piece rate.[26]

Consistent with Table 3 and Figure 7, we find performance for blueberry pickers improves by a similar amount throughout the distribution. For grapes, improvements only occur at the bottom of the distribution, with those toward the top of the distribution showing little improvement over the first 2 weeks of experience. The gap in profiles between blueberries and grapes becomes statistically significant at the 30th percentiles for standardized pieces and the latent wage, and at the median for pieces.

The evidence on learning presented thus far has used the worker-day as the unit of analysis. As such, some of our conclusions may have been driven by changes in the composition of workers at each point in the productivity distribution. In Appendix Table A1 we reproduce the analysis

[26] Specifically, the latent wage is computed by multiplying the piece rate by the number of pieces regardless of whether the minimum wage applies.

based on standardized pieces from Table 3 by analyzing learning-by-doing at the individual worker level. Reassuringly, we find similar results. Grape pickers who begin at the bottom of the productivity distribution display improvements with tenure, but the improvements for blueberry pickers are significantly steeper and exist for all quintiles of starting performance.[27]

Together, these results paint a picture consistent with the insights from our conceptual model. The commission plus bonus contract leads to considerable heaping and the bonus creates stronger incentives to learn as blueberry pickers strive to reach the threshold. Empirically, those learning incentives lead to improvements throughout the productivity distribution of blueberry workers.

4.D. Does the Bonus Pay for Itself?

As described earlier, the heaping phenomenon under the bonus contract is costly to the firm since very small changes in output around the threshold lead to sizable changes in payoff. The commission only contract does not have this feature. One way to illustrate these differences is to plot the average price the firm must pay out per piece harvested for each crop as a function of a worker's daily output, shown in Figure 10.[28]

By definition, the price per piece for blueberries jumps up discretely at 25 pieces, while for grapes the line is smooth around their piece rate threshold (32). When a blueberry picker produces 25 pieces instead of 24, the firm must pay out a $12.50 bonus. If the contract only induced this shift at the threshold, labor costs alone would dissolve almost the entire $14 revenue per piece; any other costs would likely render this extra unit of output unprofitable to the farm.

---

[27] These results are qualitatively similar when we use the other normalization approaches taken in Table 3.

[28] For clarity, the graph omits the 1% of observations where workers produce less than 8 pieces per day. Price per piece for these workers ranges from $64 for workers who pick only one piece to $9.14 for those who pick 7.

However, a different story emerges when we step away from the threshold and look at the pay schedules across the entire distribution. Importantly, price per piece steadily declines in daily output for both blueberries and grapes. The decline is steeper below the productivity target because the payout is fixed with respect to output (due to the guaranteed minimum wage). Price per piece still declines after the threshold since, although the firm pays out a piece rate, they continue to amortize the fixed wage over a larger number of pieces. The declines in price per piece at low levels of output are much larger than the discrete increase for blueberries at the productivity target. For example, at 8 pieces per day, the firm must pay out $8 per piece, which is more than half the revenue from selling each piece. If a worker improves from 8 pieces per day to 12 (4 pieces is the typical improvement over 10 days at most points in the distribution, see figure 8), the firm saves $3 on each piece produced, which is one-quarter of the revenue the firm receives for it. In contrast, a move from 24 to 25 pieces increases the average price per piece by only $0.40.

The firm thus makes large revenue gains if their contract induces low-end workers to improve their performance. The previous section highlighted the fact that blueberry workers under the bonus contract exhibit considerable learning on the job throughout the productivity distribution. As such, one reason firms might offer a bonus contract that leads to costly heaping is that it also leads to beneficial improvements by increasing the productivity of the costliest and least productive workers. The question is whether the productivity improvements at the low end dominate the inefficient heaping just above the piece rate threshold.

To assess this trade-off, we perform a simple back of the envelope calculation. As a benchmark, we begin with grapes. Comparing workers with less than 4 days on the job to those with 10 days, we find that average output increases by 1.4 pieces (from 31.4 to 32.8). At the same time, the firm must pay out $0.30 to $0.35 for each piece picked above the threshold. We already know this cost is small – reaching the threshold is minimally related to experience (table 2) and that productivity does not improve by much for workers above the threshold (table 3 and figure 9). Empirically, this cost change averages only $0.05. On average, the firm gets an additional 1.4

pieces per worker at a cost of $0.05. When factoring in the price per piece that the firm can expect to receive for its output ($0.33 per pound at 5 pounds per piece), we find that profits increase by $2.26 per worker for a worker with 10 days of experience relative to one in his first 3 days.

For blueberries the calculation differs on both sides: production increases by substantially more, as do costs. Comparing workers with less than 4 days on the job to those with 10 days, average output increases by nearly 5 pieces (from 16.4 to 21.2). At the same time, the extra 10 days of experience appreciably increases a blueberry picker's chance of surpassing the threshold (by 24 percentage points), requiring the firm to pay out $12 dollars when this happens, plus $0.50 on each additional piece picked above the threshold. Empirically, these improvements increase costs by $3.40 per worker, on average. Thus, the firm gets an average of nearly 5 additional pieces for an average cost of $3.40 per worker. Factoring in the larger revenue the firm receives per blueberry bucket, the firm can expect a little over $60 per worker in profits for blueberries for a worker with 10 days of experience relative to one with less than 4, which is approximately $9 per worker per day. Thus, it appears that the improvement in output induced by the bonus contract is extremely profitable for the firm, even though the bonus payment is 'locally' distortionary at the productivity threshold.

5. Additional Results

In this section we probe the mechanisms through which workers improve their productivity and examine alternative explanations for our results.

5A. Improvements in picking speed or avoiding burn-out?

To explore possible mechanisms through which blueberry workers improve their productivity, we utilize the subset of workers for whom we have hourly productivity data to investigate two possible stories. One possibility is that more experienced workers become more competent at the physical act of picking blueberries. In this case, we expect productivity for more experienced workers to be higher throughout the day. An alternate explanation is that more experienced workers become better at pacing themselves. In this case, we expect output of the less experienced workers to decrease towards the end of the day.

In Figure 11, we plot hourly productivity for blueberry harvesters by tenure profiles in roughly 3-day bins (1-2, 3-5, 6-8, 9-10, and 10+). The figure shows output in each hour of the day from 6am to 2pm. We find that productivity at all hours of the day increases with tenure. For those with 10 or more days of tenure, productivity is nearly 50% higher at all hours of the day compared to those with less than 3 days. The other workers lie in between these groups. We find little evidence of fatigue for any tenure group, as output does not decline throughout the day.[29] This suggests workers become more competent at picking blueberries with more experience.

5B. Alternative Explanations

Our empirical work takes advantage of the different contract regimes across crops. We find substantially more learning-by-doing for blueberry pickers than for grape pickers and attribute this to the different contract regimes. However, there may be other differences across the two crops that generate these results.

First, perhaps less scope for learning exists in grapes compared to blueberries. While we cannot rule this out entirely, we note that Tables 2 and 3 show some improvements with tenure for grape pickers, with figures 8 and 9 showing these improvements coming from the bottom of the performance distribution. Also, figure 4 shows substantial heterogeneity in grape picker performance, with the best producers out-picking the worst by a factor of 3-4, comparable to the variation in blueberries. Thus, there appears to be scope for productivity improvements in grape harvesting.

[29] In fact, we see some evidence of an uptick in productivity in the last hour of the day. This is likely driven by a late surge in effort for workers close to the threshold.

Second, one may worry about another form of gaming: workers sharing their pickings with each other to maximize group payoffs. We cannot fully rule out this explanation, but oversight by field supervisors is meant to limit this sort of behavior. To the extent that it could occur, it is generally only incentive compatible between workers who have already crossed the productivity target and those just shy of it. This would amplify heaping but is unlikely to explain productivity shifts at other parts of the distribution. If a worker above the productivity target offers a piece to a worker who does not reach the target, she is losing the piece rate the farm would have given her. Meanwhile, the worker who accepts the piece remains in the fixed wage regime and is not receiving additional compensation. There may be other possible gaming opportunities in the field, but they would have to take a very specific form to explain our finding that output improves at all parts of the performance distribution for blueberry but not for grape pickers.

Finally, workers may non-randomly sort into crops such that those least able to succeed with the commission plus bonus contract select into grapes instead of blueberries. To explore this hypothesis, we take advantage of the fact that we observe 7% of our workers in both grapes and blueberries. A worker in the spring/early summer blueberry season might stay on for the late summer/fall grape harvest, or a grape worker may return for blueberries the following year. Among workers in multiple crops, two-thirds start out in blueberries and one-third start out in grapes, which is not surprising given the longer time lag from grapes to blueberries.

Appendix figure A1 shows productivity histograms for workers in both crops, separately by current crop and tenure. These histograms look quite similar to those previously presented. In the top panel, we see little heaping for blueberry pickers with low tenure (top) but considerable heaping for those with high tenure (bottom). In the right panels, we see that neither low nor high tenure grape pickers heap. That those with prior experience in another crop perform similarly to those native to that crop suggests that worker selection is not driving our results.

6. Conclusion

Pay-for-performance compensation schemes are used across a wide range of industries and occupations. Even contracts that offer a fixed salary may be implicitly tied to performance by promotion and termination incentives (Lazear, 2000). Firms offer performance-based pay to motivate workers to supply costly effort, but they may accomplish more than simply inducing employees to work harder in applying their existing skill set. A large body of literature has focused on undesirable outcomes induced by performance-based pay, such as gaming. In this paper we propose that such contracts may also induce workers to acquire new skills that make them better at their job, a much more desirable outcome from the firm's perspective.

We exploit a unique panel dataset on individual worker productivity to provide the first empirical evidence on the impacts of performance-based compensation schemes, and particularly the use of bonus payments, for on-the-job learning. We find that contracts with bonuses tied to productivity targets generate faster learning-by-doing than those induced by a simple commission scheme. Moreover, these productivity improvements are not simply limited to those close to the bonus threshold. Instead, learning occurs throughout the productivity distribution. We find that the costly heaping induced by the highly nonlinear bonus structure is more than offset by the overall gains in productivity from those well below the bonus threshold.

This set of results is important for several reasons. Firms collectively spend billions of dollars each year on research and development in a quest for technologies that can help improve firm productivity and profitability. At the same time, they also invest considerable sums in human resource management to both monitor and train workers to better utilize these new, as well as older, technologies. In our setting, compensation schemes designed to incentivize effort also play an important role in learning and skill acquisition by workers. This suggests such incentives may be an important tool to add to the armamentarium of instruments used to improve firm productivity across many different settings.

The dynamic implications of learning-by-doing also raise new and interesting questions about the design of contracts aimed at addressing principal-agent problems. While performance-based pay can induce workers to supply costly effort, it can also affect learning. After all, if a skill learned is not easily forgotten (Friedman and Kelman, 2007), and skills are quite transferrable across occupations, then moral hazard in effort is likely to be much stronger than in learning. On the other hand, if learning-by-doing is itself quite costly and slows down output, depreciates quickly, or is firm-specific, then the opposite may be true. Understanding the conditions under which the incentives to induce learning-by-doing are complements or substitutes to those needed to induce imperfectly observed effort is fundamental for optimal contract design.

While our results point to the importance of bonus payments and performance targets for learning on the job, there are many unanswered questions. How important is the salience of these goals versus the actual monetary payoff for incentivizing workers to become increasingly skilled in their jobs? What are the criteria shaping optimal bonus target setting? A low target will only provide low-powered incentives for learning and may lead to too much inefficient heaping, while a high target may also induce little learning should most workers feel the bonus is beyond their reach. Given the ubiquity of bonus contracts across sectors that include sales, corporate executives, and even physicians, how might these results generalize? Together, they comprise a future research agenda that can create stronger connections between two fundamental but under-linked literatures in economics.

7. References

Abowd, J. and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," Econometrica, 57(2): 411-455.

Altonji, J.G. and C.R. Pierret (2001): "Employer Learning and Statistical Discrimination," Quarterly Journal of Economics, 113:79-119.

Argote, Linda, Sara L. Beckman, and Dennis Epple. 1990. "The Persistence and Transfer of Learning in Industrial Settings." Management Science 36(2): 140–54.

Arrow, Kenneth J. 1962. "The Economic Implications of Learning by Doing." Review of Economic Studies 29(3): 155–73.

Baker, George, 1992. "Incentive Measures and Performance Measurement," Journal of Political Economy, 100, 598-614.

Baker, M. (1997): "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," Journal of Labor Economics, 15: 338-375.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, 2011. "Field Experiments with Firms," Journal of Economic Perspectives, 25(3), 63-82.

Becker, G.S., 1962. Investment in human capital: A theoretical analysis. Journal of political economy, 70(5, Part 2), pp.9-49

Ben-Porath, Y. (1967), "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, vol LXXV: pp. 352-365.

Courty, Pascal; and Gerald Marschke, 2004. "An Empirical Investigation of Gaming Responses to Explicit Performance Incentives," Journal of Labor Economics, 22(1), 23-56.

Farber, H.S. and R. Gibbons (1996): "Learning and Wage Dynamics," Quarterly Journal of Economics, 111:1007-1047.

Frandsen, Brigham R. (2017), Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete, in Matias D. Cattaneo , Juan Carlos Escanciano (ed.) Regression Discontinuity Designs (Advances in Econometrics, Volume 38) Emerald Publishing Limited, pp.281 - 315

Freeman, Richard B., Wei Huang, and Teng Li, "Non-Linear Incentives and Worker Productivity and Earnings: Evidence from a Quasi-Experiment," NBER Working Paper #25507 (2019).

Friedman, John N., and Steven Kelman. "Effort as investment: analyzing the response to incentives." (2007).

Gibbons, Robert, and Kevin J. Murphy. "Optimal incentive contracts in the presence of career concerns: Theory and evidence." Journal of political Economy 100.3 (1992): 468-505.

Graff Zivin, Joshua, and Matthew Neidell. “The Impact of Polution on Worker Productivity.” American Economic Review 102 (December 2012): 3652-3673.

Grossman, Sanford J., and Hart, Oliver D. "An Analysis of the Principal-Agent Problem." Econometrica 51 (January 1983): 7-45.

Guvenen, Fatih (2007): "Learning your Earning: Are Labor Income Shocks Really Very Persistent?" American Economic Review, 97(3): 687-712.

Harris, M., AND A. Raviv, "Optimal Incentive Contracts with Imperfect Information," Journal of Economic Theory, 20(1979), 231-259.

Hause, J. (1980): "The Fine Structure of Earnings and the On-the-Job Training Hypothesis," Econometrica, 48: 1013-29.

Healy, Paul M. "The effect of bonus schemes on accounting decisions." *Journal of accounting and economics* 7.1-3 (1985): 85-107.

Holmstrom, B., "Moral Hazard and Observability," Bell Journal of Economics, 10(1979), 74-91.

Holmstrom, Bengt, and Paul Milgrom. "The firm as an incentive system." The American Economic Review (1994): 972-991.

Jovanovic, Boyan, and Yaw Nyarko. "A Bayesian learning model fitted to a variety of empirical learning curves." Brookings Papers on Economic Activity. Microeconomics 1995 (1995): 247-305.

Kahn, Lisa B. and Fabian Lange (2014), “Employer Learning, Productivity and the Earnings Distribution: Evidence from Performance Measures,” *Review of Economic Studies*, 81(4): pp. 1575-1613.

Kleven, Henrik Jacobsen, et al. "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark." *Econometrica* 79.3 (2011): 651-692.

Lange, F. (2007): "The Speed of Employer Learning." Journal of Labor Economics, 25: 1-35.

Larkin, Ian "The Cost of High-Powered Incentives: Employee Gaming in Enterprise Software Sales," Journal of Labor Economics 32, no. 2 (April 2014): 199-227.

Lazear, Edward P., 1986. “Salaries and Piece Rates,” Journal of Business, 59(3), 405-431.

Lazear, Edward P., 2000, “Performance Pay and Productivity,” American Economic Review, 90(5), 1346-1361.

Levitt, Steven D., John A. List, and Chad Syverson. "Toward an understanding of learning by doing: Evidence from an automobile assembly plant." Journal of Political Economy 121.4 (2013): 643-681.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. Psychological Bulletin, 90(1), 125-152.

Locke, Edwin A., and Gary P. Latham. 1990. A theory of goal setting and task performance. Englewood Cliffs, NJ: Prentice Hall.

MaCurdy, T.(1982) "The Use of Time Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis," Journal of Econometrics, 18: 83-114.

McCrary, Justin (2008). Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics, 142:698–714.

Mincer, J. (1958) *Schooling, Experience, and Earnings,* New York, NY, National Bureau of Economic Research.

Mirrlees, J., "The Optimal Structure of Incentives and Authority Within an Organization," Bell Journal of Economics, 7(1976), 105-131.

Oyer, Paul, 1998. “Fiscal Year Ends and Non-Linear Incentive Contracts: The Effect on Business Seasonality,” Quarterly Journal of Economics, 113, 149-185.

Prendergast, Canice, 1999, “The Provision of Incentives in Firms,” Journal of Economic Literature, 37(1), 7-63.

Prendergast, Canice, 2002, “The Tenuous Trade-Off between Risk and Incentives,” Journal of Political Economy, 110(5), 1071-1102.

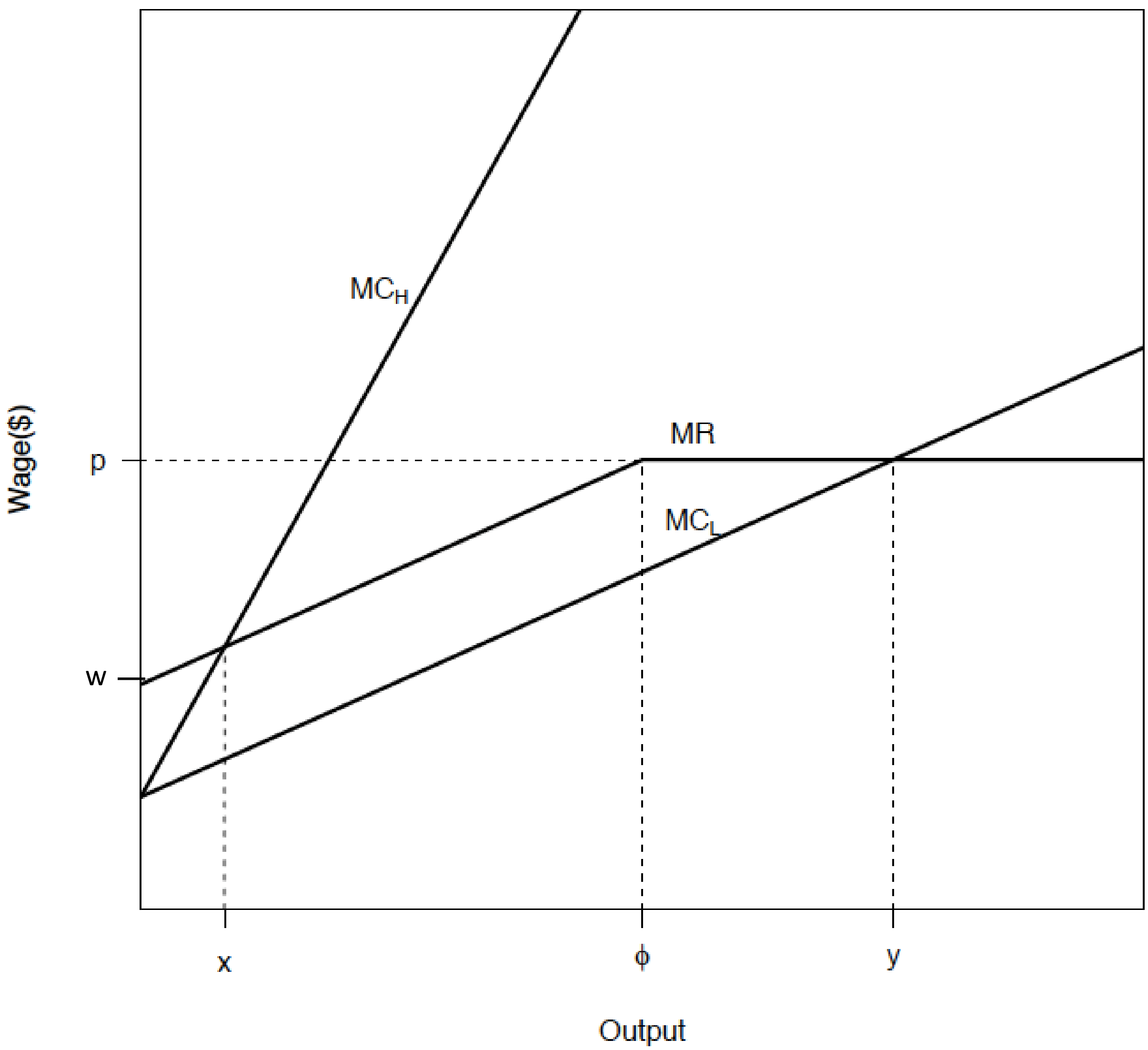
Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic perspectives*, *8*(1), 3-22.

Saez, Emmanuel. "Do taxpayers bunch at kink points?." *American Economic Journal: Economic Policy* 2.3 (2010): 180-212.

Stevens, Andrew. “Temperature, Wages, and Agricultural Labor Productivity.” (2018) Working paper.

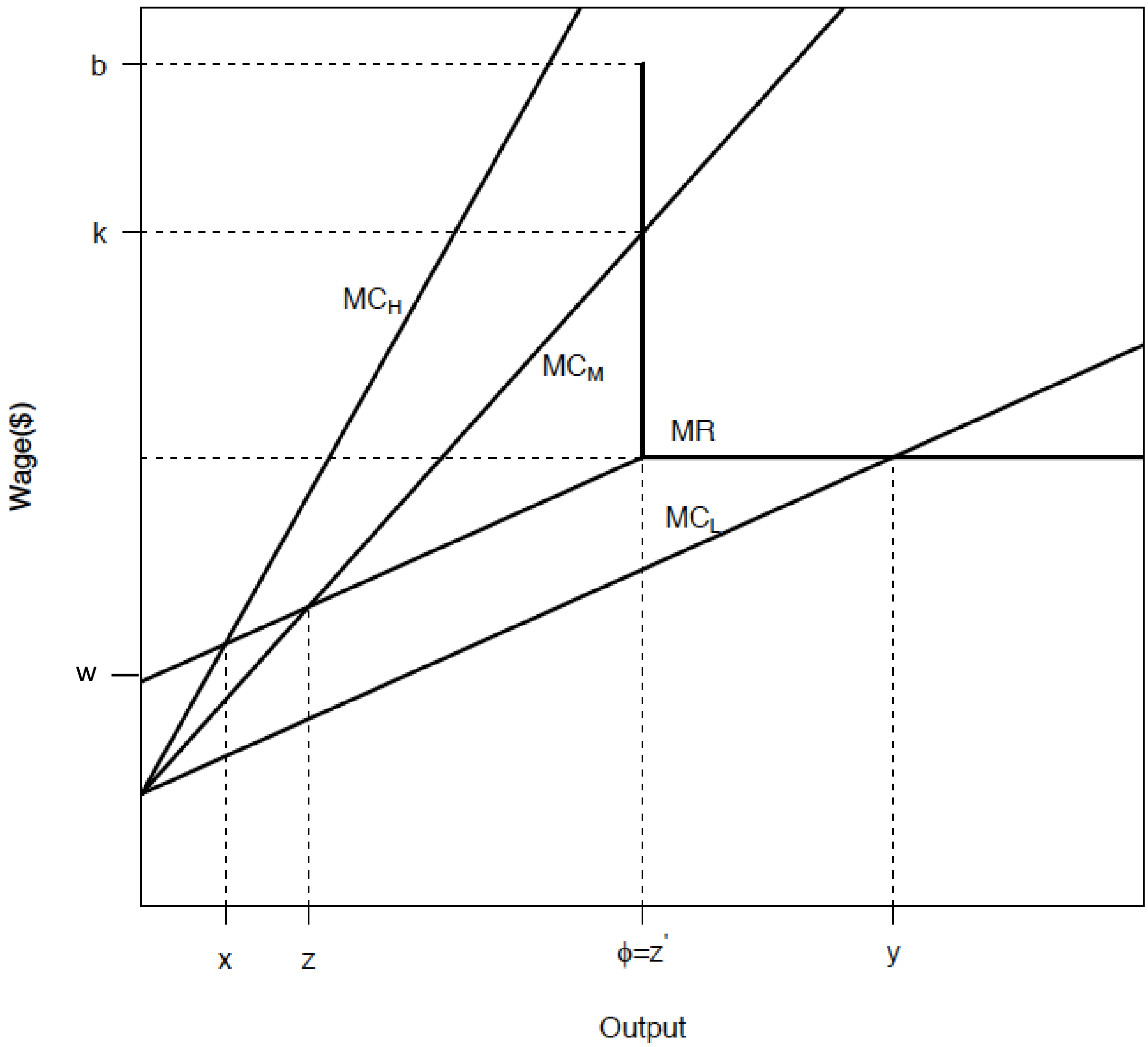
Thompson, Peter. "The relationship between unit cost and cumulative quantity and the evidence for organizational learning-by-doing." The Journal of Economic Perspectives 26.3 (2012): 203-224.

Figure 1: Output and Wages under Commission Contract



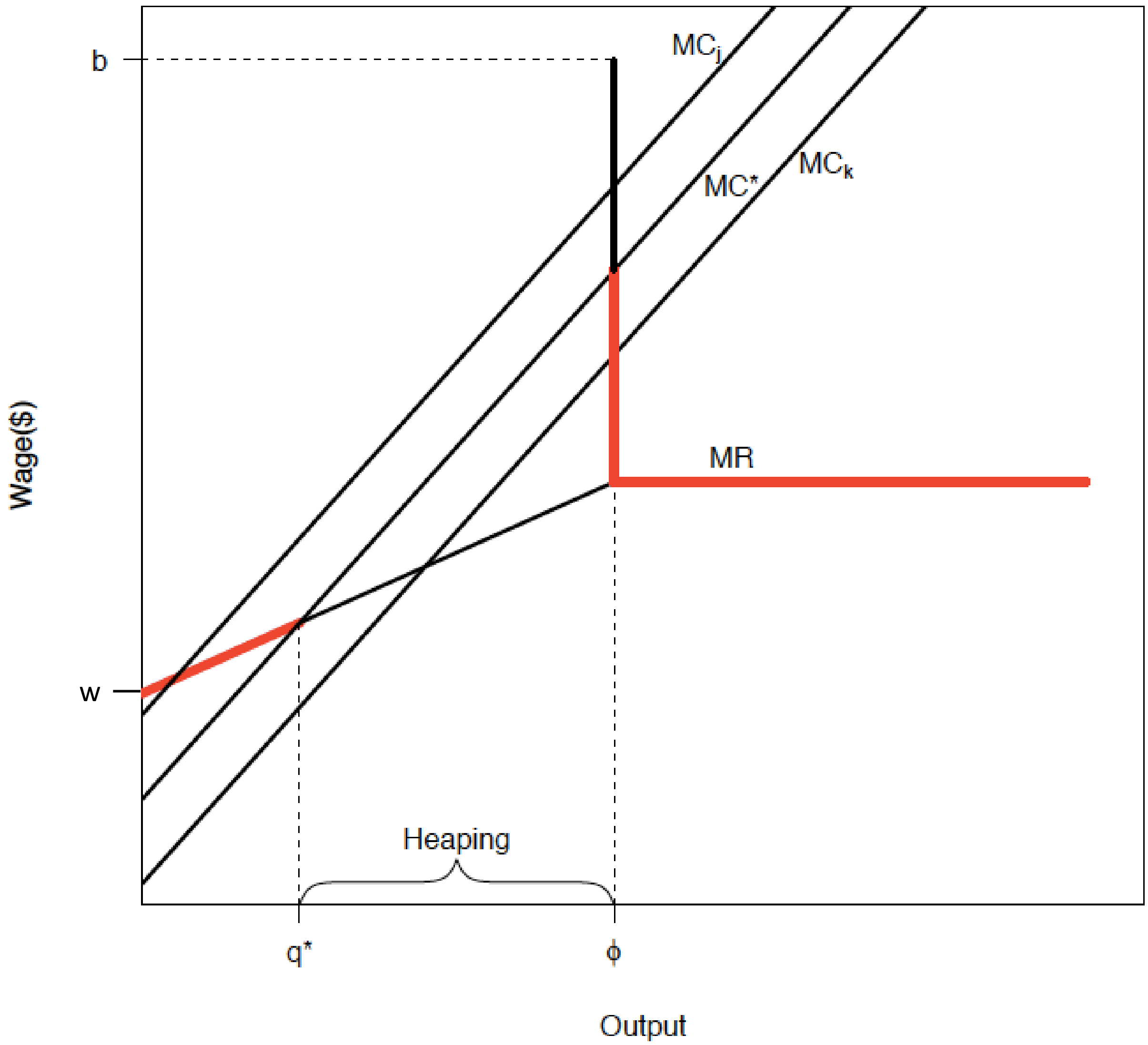
Notes: MR is the marginal take home revenue (a function of termination probability adjusted earnings) curve for the worker. The target productivity level is denoted ϕ. MC$_H$ is the marginal cost curve for a high effort cost type worker who produces output x. MC$_L$ is the marginal cost curve for a low effort cost type worker who produces output y. The MR curve has a value of w at zero output, the minimum guaranteed wage. It is then upward sloping until point ϕ, the productivity target, to reflect the fact that increases in output reduce the worker's termination probability. At point ϕ, the worker receives piece rate, p, per unit output and suffers no firing risk (by assumption).

Figure 2: Output and Wages under Commission plus Bonus Contract



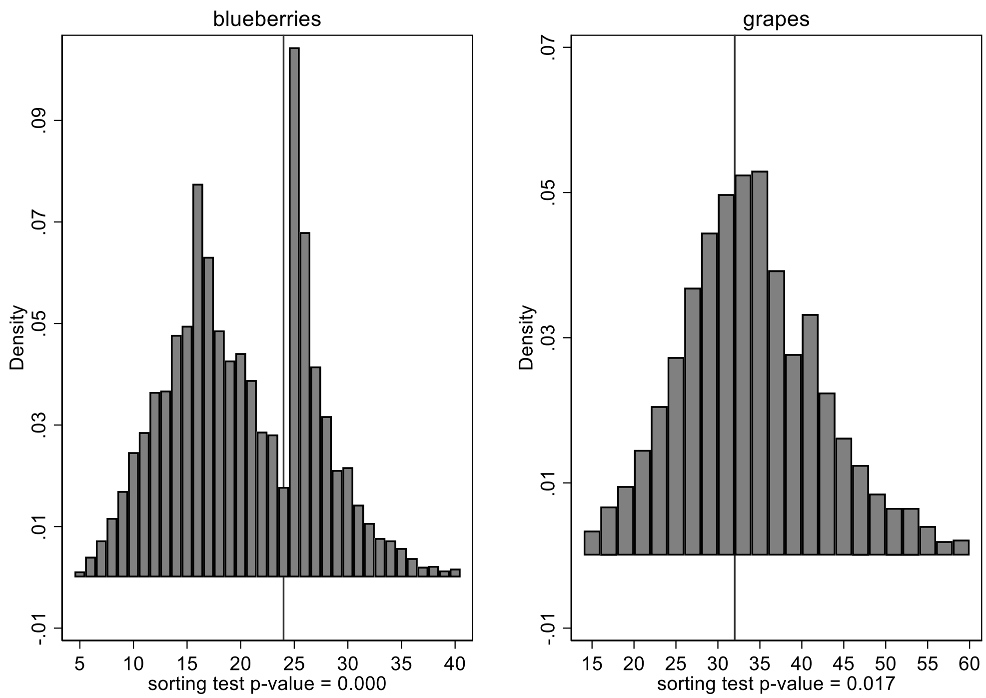
Notes: See figure 1. MR is the marginal revenue (termination probability adjusted earnings) curve for the worker, which includes a spike at the target productivity level ϕ up to the bonus payment b. MC_H is the marginal cost curve for a high effort cost type worker, who produces x units of output. MC_L is the marginal cost curve for a low effort cost type worker, who produces y units of output. MC_M is the marginal cost curve for a moderate effort cost type worker, who will either produce z or z' units of output. This worker produces z' if $\int_z^{\phi}[c'(q) - f'(q)]\,dq \leq [\mathrm{b} - k]$ where k is the marginal cost of producing z' for the moderate type worker.

Figure 3: Heaping under Commission plus Bonus Contract



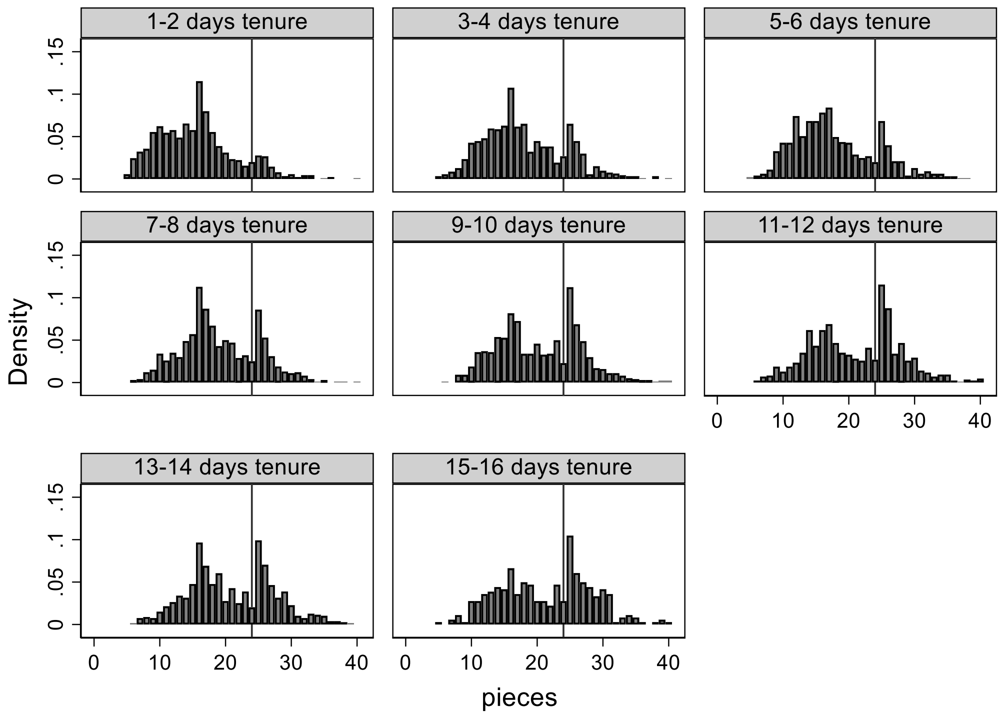
Notes: MR is the marginal revenue (termination probability adjusted earnings) curve for the worker, which includes a spike at the target productivity level ϕ up to the bonus payment b. MC^* is the marginal cost curve that intersects the MR curve precisely at q*, such that the marginal gains from hitting the production target ϕ are precisely equal to the excess costs of reaching this target. Worker i with MC_i will produce output below ϕ and he will be paid the minimum wage, w. Worker k with MC_k will find it worthwhile to reach the productivity target and thus produce ϕ units of output. Heaping will occur for all moderate effort cost type workers with a marginal cost curve that intersects the upward sloping segment of the MR curve between q* and ϕ. Equilibrium output across all worker types is highlighted in red and bold.

Figure 4. Productivity distribution by crop



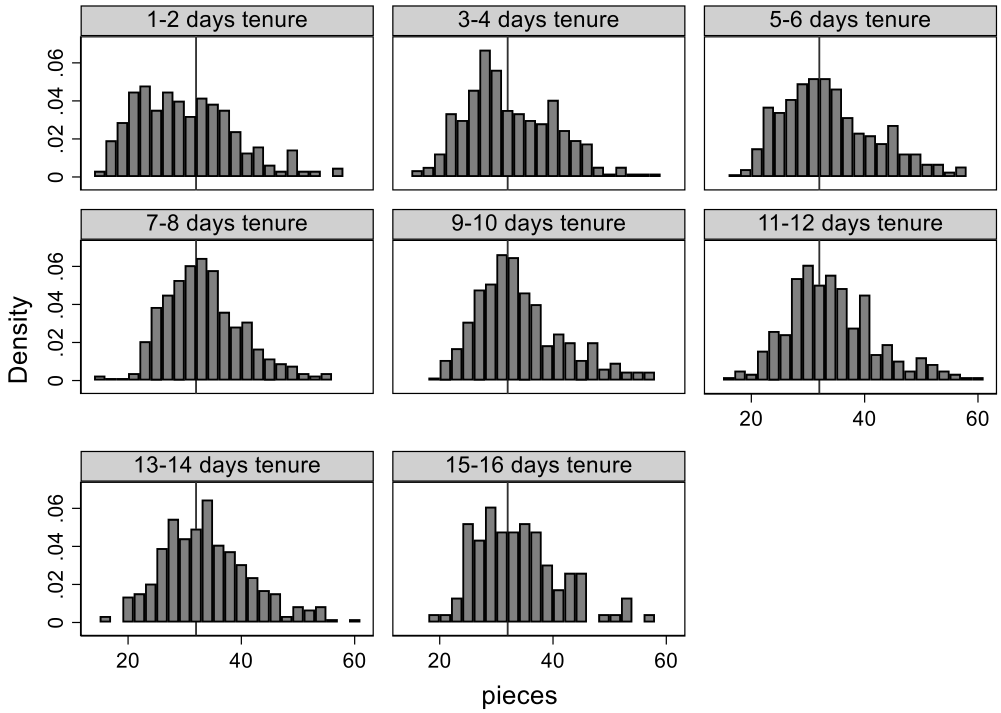
Notes: We plot histograms of daily output (pieces), separately by crop, for workers that eventually complete at least 10 days. The vertical lines show the threshold at which workers cross from the minimum wage to the piece rate regime. We report p-values for the Fransden (2017) heaping test, testing whether the performance distribution is smooth across the piece rate threshold.

Figure 5. Blueberry productivity by tenure



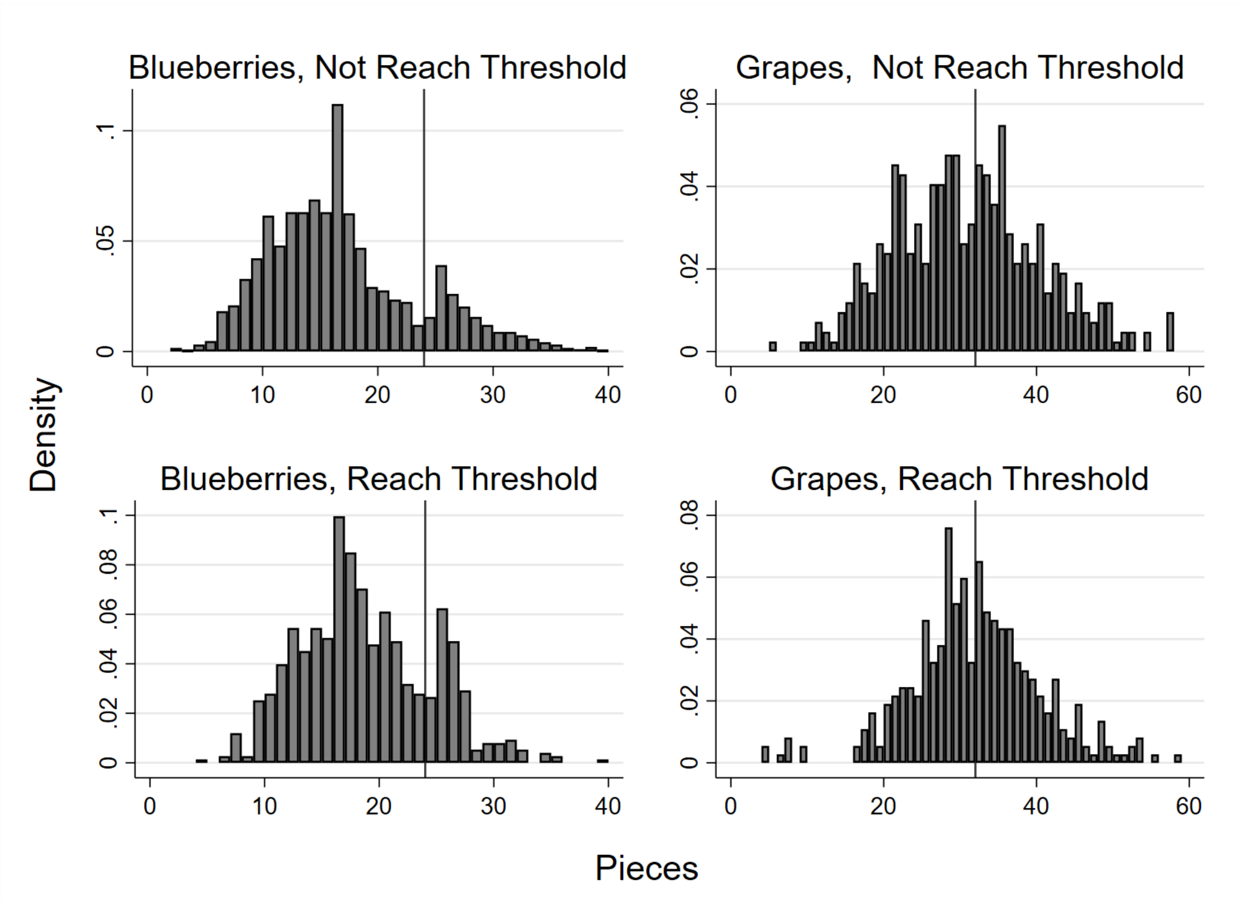
Notes: Graphs are restricted to blueberry pickers. We plot histograms of daily output (pieces) for the indicated number of days tenure. All graphs restrict the sample to workers who eventually survive at least 10 days so the composition of workers is essentially fixed across graphs (through the 9-10 days graph). The vertical lines show the threshold at which workers cross from the minimum wage to the piece rate regime.

Figure 6. Grape productivity by tenure



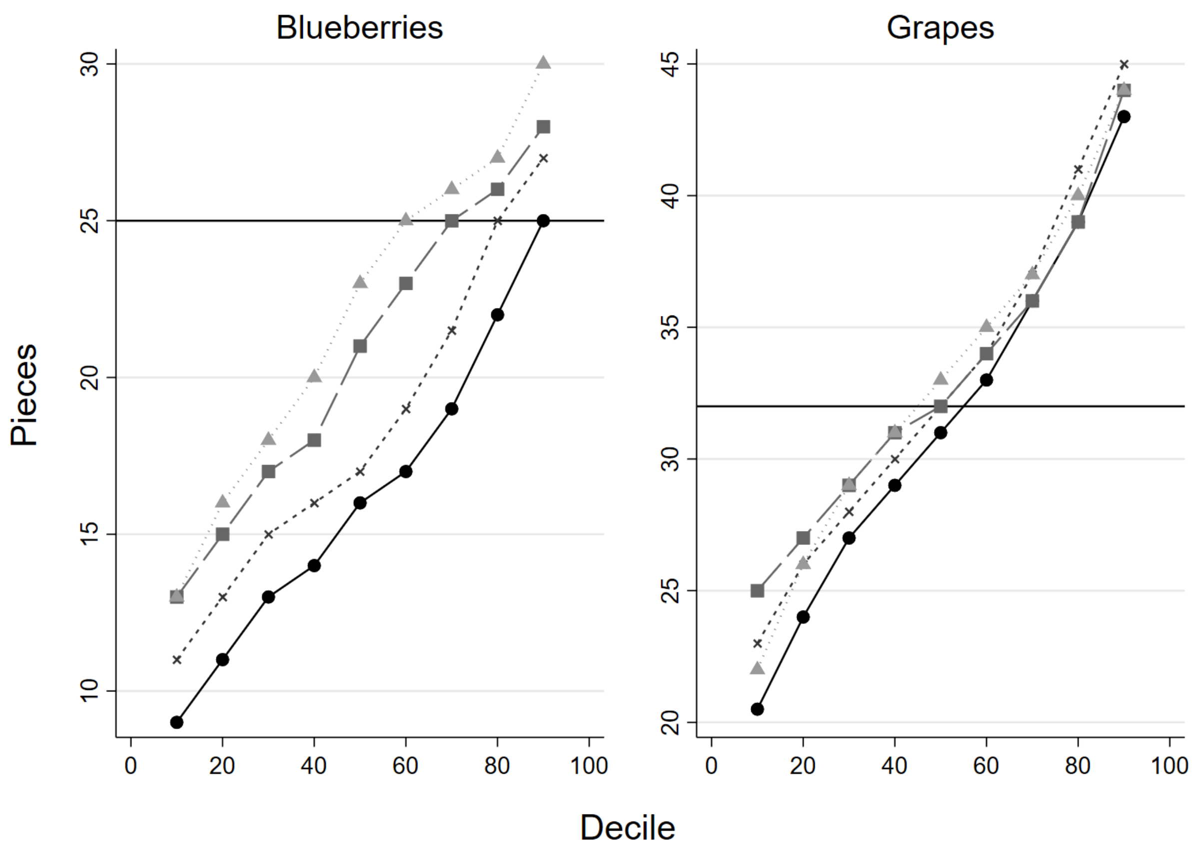
Notes: See figure 5. All graphs are restricted to grape pickers who eventually survive at least 10 days.

Figure 7. Day 1-3 Performance by whether the worker hits the piece rate threshold on days 10-12



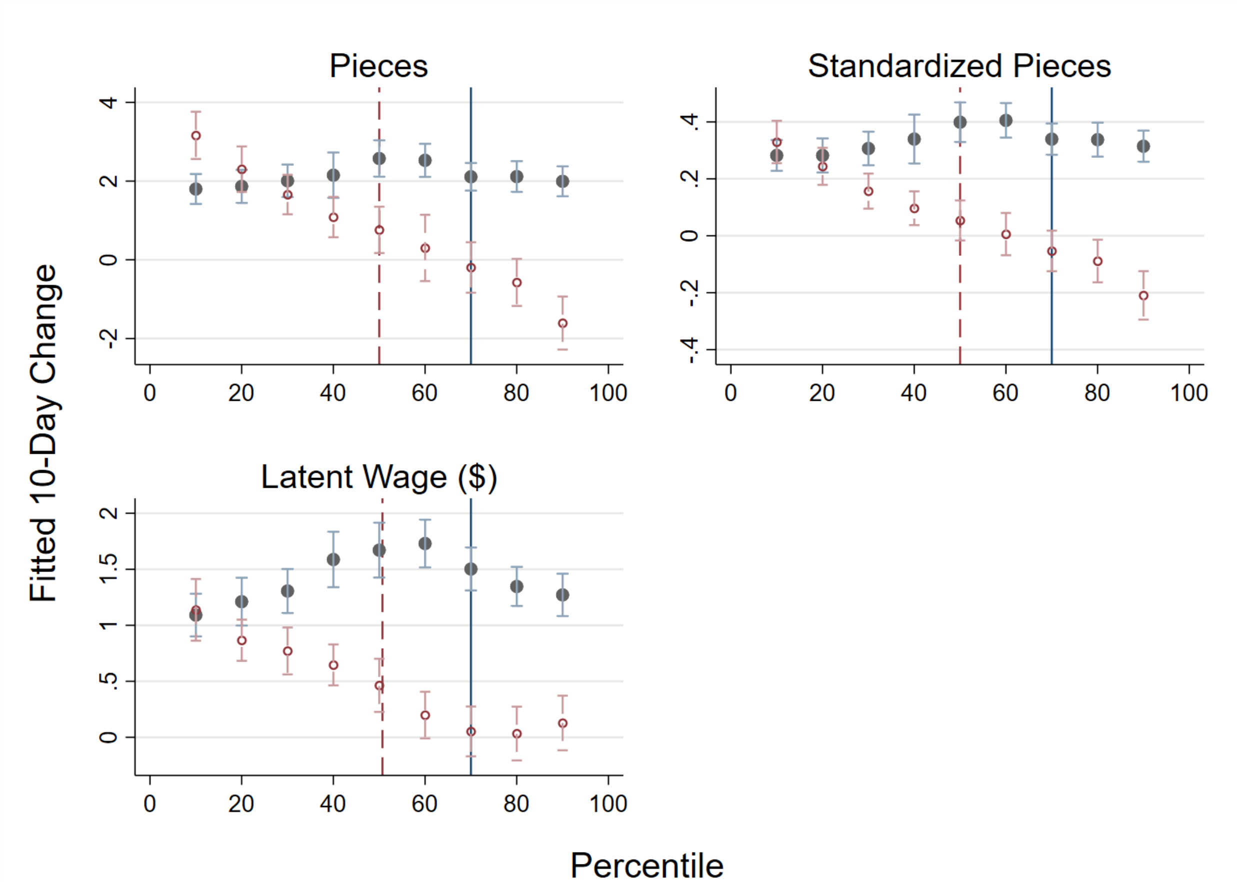
Notes: We plot histograms of day 1-3 performance for workers who eventually survive at least 10 days. The top panel shows the distributions, by crop, for workers who do not reach the piece rate threshold on days 10-12. The bottom panel shows the distribution for workers who reach the piece rate threshold at least once on days 10-12.

Figure 8: Performance Decile by Tenure and Crop



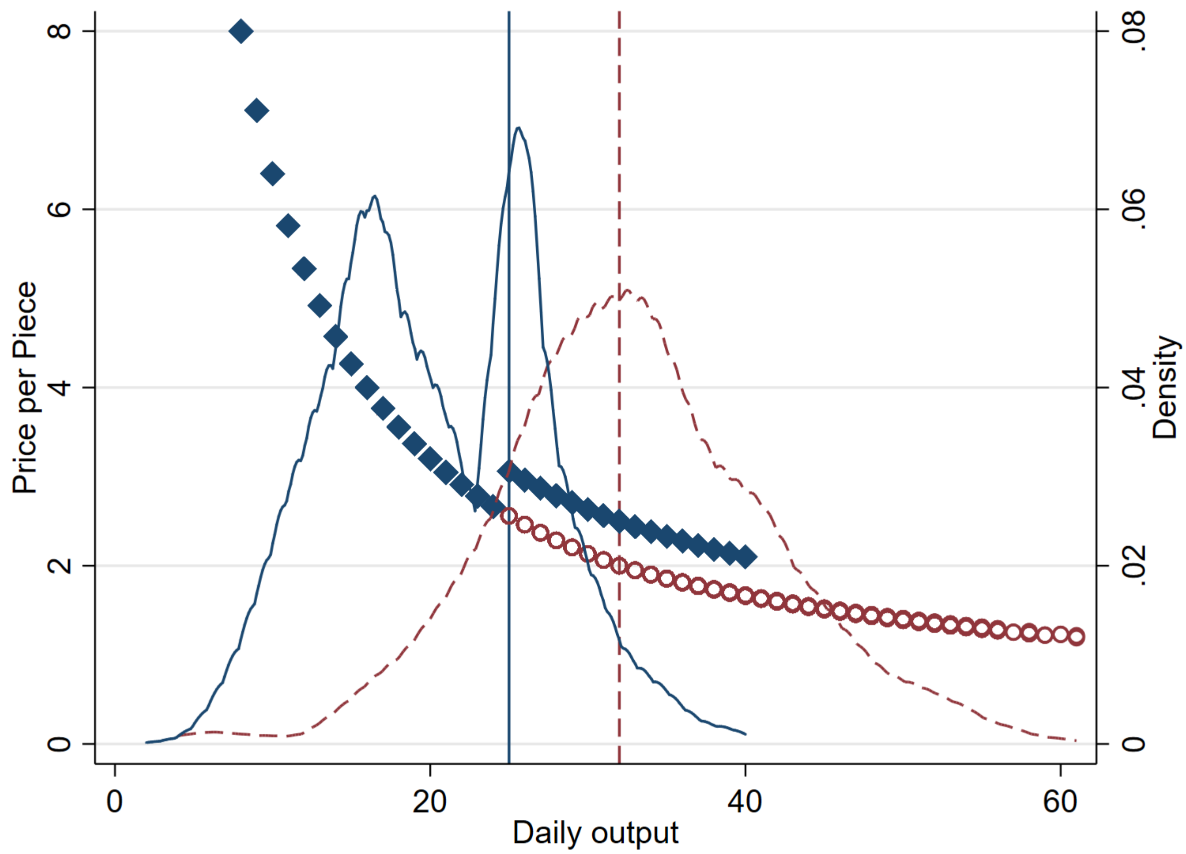
Notes: Darker colors indicate newer workers. Circle = tenure < 4 days, X = tenure 4-6 days, square = tenure 7-9 days, and triangle = tenure 10 or more days. Figure plots the daily output associated with each decile in the distribution of pickers by crop and tenure group. The sample is restricted to workers whose tenure at the farm lasts at least 10 days and to workers in their first crop. The piece rate threshold for blueberry (grape) pickers is 25 (32), indicated with a horizontal line.

Figure 9. Fitted 10-Day Performance Growth, by Decile, Tenure and Crop



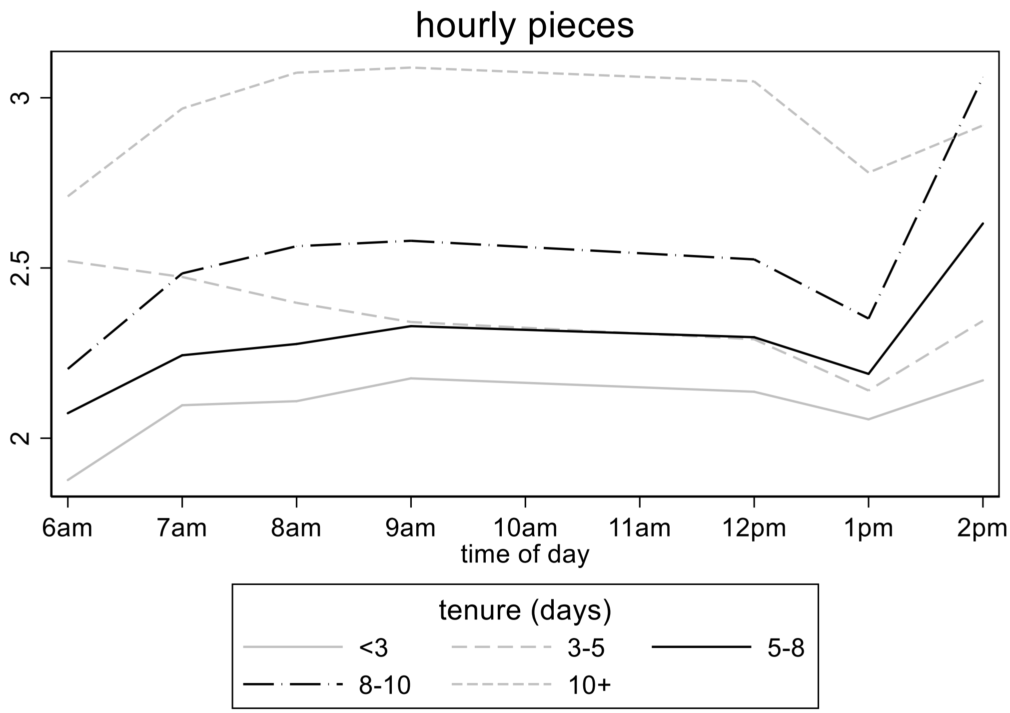
Notes: Solid = blueberries, hollow = grapes. Capped vertical lines indicate 90% confidence intervals. Using quantile regression, we estimate quadratic tenure profiles interacted with crop, for each performance decile, controlling for gender, day-of-week, week-of-harvest, and first/last harvest day. We plot fitted effects for the 10th day along with 90% confidence intervals. We restrict to workers who survive at least 10 days and are in their first crop. Vertical lines indicate the point in the distribution where the piece rate thresholds are for each crop (blueberry line is solid, grape line is dashed). Standardized pieces (top right) are pieces normed to be mean 0, standard deviation 1, within crop. Latent Wage (bottom left) are the number of pieces times the piece rate ($0.50 for blueberries, $0.30 or $0.35 for grapes).

Figure 10. Price per Piece and Density of Daily Output



Notes: Solid/diamonds = blueberries; hollow/dashed = grapes. Figure plots the relationship between daily price per piece (assuming an 8-hour workday) and output (dots), as well as the realized output distribution (lines), by crop. For clarity, the price per piece dots omit output<8 pieces (1% of the sample); price per piece for these workers ranges from $64 for workers who pick only one piece to $9.14 for those who pick 7. Vertical lines indicate the piece rate threshold for each crop, 25 for blueberries (solid line) and 32 for grapes (dashed). Histograms in the figure are restricted to workers who survive at least 10 days and are in their first crop.

Figure 11. Hourly productivity for blueberries by tenure



Notes: Figure plots hourly productivity for blueberry pickers by tenure group (roughly 3-day bins indicated in the legend). We restrict the sample to workers who eventually survive at least 10 days. The figure shows output in each hour of the day from 6am to 2pm.

Table 1. Summary Statistics

	Blueberries	Grapes
Pieces	20.38	32.81
	(6.74)	(8.75)
Daily Earnings ($)	68.99	65.20
	(6.73)	(1.77)
Hits piece rate threshold	0.36	0.55
	(0.48)	(0.50)
Tenure (days)	12.00	13.74
	(9.15)	(11.49)
Completed Tenure	24.33	28.41
	(11.63)	(14.30)
Survives ≥ 10 Days	1	1
	(0)	(0)
# worker-day obs	9,953	3,980
# unique days	57	44
# unique workers	555	244

Notes: Table reports means (and standard deviations) for workers who eventually survive at least 10 days. Pieces is the number of blueberry buckets or grape boxes (as indicated) picked in the day. The piece rate threshold is 32 for blueberries and 25 for grapes. Daily earnings normalizes to an 8-hour work day.

Table 2. Probability of exceeding piece rate threshold as function of tenure and crop

Dependent Variable: Hit Piece Rate Threshold				
	(1)	(2)	(3)	(4)
Blueberries	-0.309***	-0.400***	-0.401***	-0.350***
	(0.029)	(0.032)	(0.032)	(0.033)
Tenure/10	0.042***	0.035	0.035	0.057*
	(0.010)	(0.027)	(0.027)	(0.031)
$Tenure^2/1000$		0.014	0.014	-0.003
		(0.053)	(0.052)	(0.060)
Blueberries*Tenure/10	0.105***	0.265***	0.266***	0.105***
	(0.016)	(0.034)	(0.033)	(0.037)
Blueberries*$Tenure^2/1000$		-0.445***	-0.444***	-0.198**
		(0.077)	(0.076)	(0.083)
Constant	0.493***	0.497***	0.516***	0.472***
	(0.024)	(0.028)	(0.034)	(0.035)
Observations	13,933	13,933	13,933	13,933
R-squared	0.086	0.094	0.095	0.174
Gender Control			X	X
Date Controls				X

Notes: *** p<0.01, ** p<0.05, * p<0.1; standard errors are clustered by worker. We regress an indicator for whether the worker hit his piece rate threshold on the given day. The threshold for grapes is 32 pieces and the threshold for blueberries is 25. The sample is restricted to workers who are in their first crop and eventually survive for at least 10 days. Gender controls are indicators for male and female (with missing gender as the omitted category). Date controls include day of week, week of harvest, and first and last day of harvest indicators.

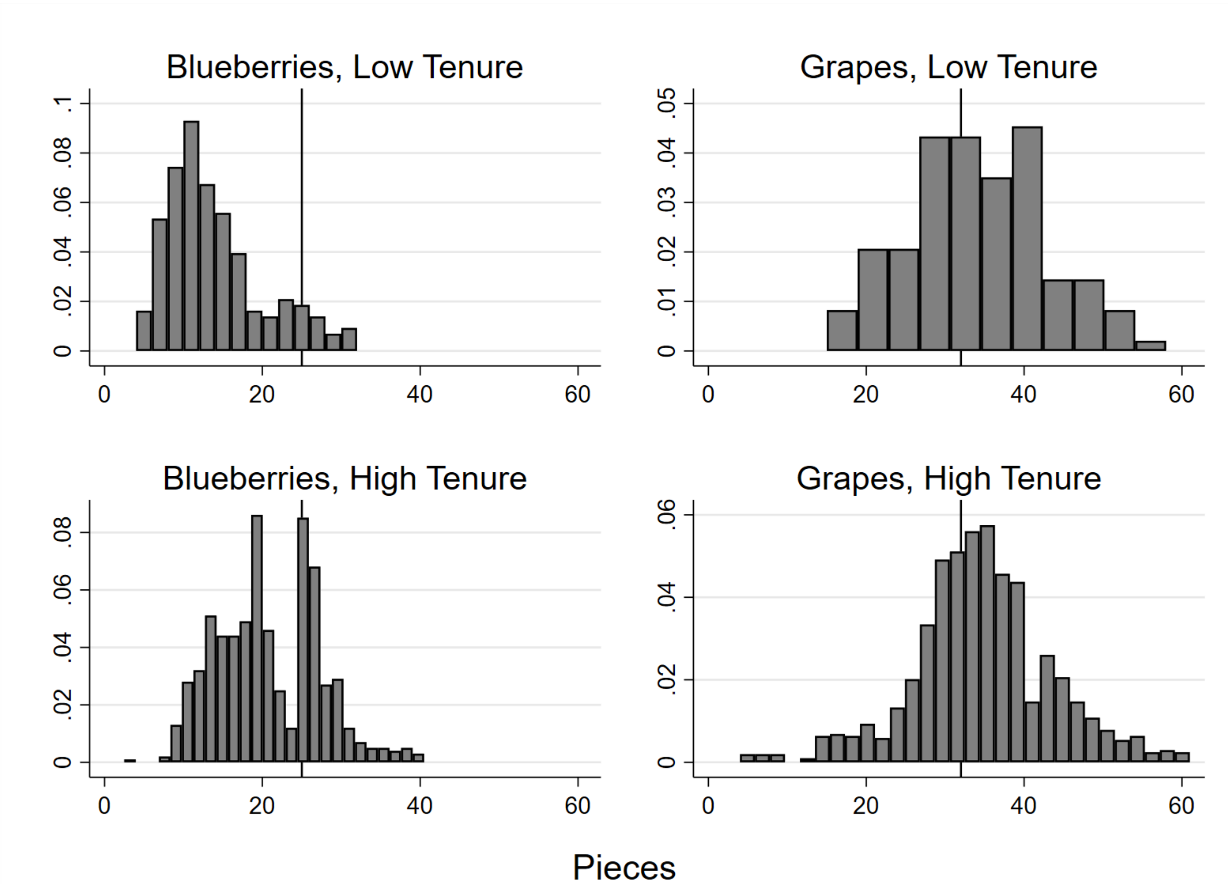
Table 3: Productivity Tenure Profiles throughout the Productivity Distribution

Dependent Variable: Daily Output Quintile for Standardized Pieces				
Quintile:	20	40	60	80
Blueberries	-0.190**	-0.359***	-0.380***	-0.526***
	(0.093)	(0.098)	(0.098)	(0.124)
Tenure/10	0.274***	0.087	-0.026	-0.130
	(0.077)	(0.077)	(0.103)	(0.101)
Tenure^2/1000	-0.299*	0.094	0.319	0.409**
	(0.168)	(0.152)	(0.227)	(0.187)
Blueberries*Tenure/10	0.038	0.288**	0.511***	0.528***
	(0.096)	(0.124)	(0.107)	(0.112)
Blueberries*Tenure^2/1000	0.000	-0.452	-1.111***	-1.016***
	(0.208)	(0.322)	(0.243)	(0.217)
Constant	-1.141***	-0.539***	-0.160	0.557***
	(0.099)	(0.099)	(0.106)	(0.119)
Observations	13,933	13,933	13,933	13,933
R-squared	0.155	0.183	0.175	0.169

Notes: *** p<0.01, ** p<0.05, * p<0.1; standard errors are clustered by worker. We estimate quantile regressions of standardized daily performance (mean 0, standard deviation 1 within crop) for each indicated quantile. Regressions include controls for gender and date (see table 2). The sample is restricted to workers who are in their first crop and eventually survive for at least 10 days.

Appendix

Appendix Figure A1. Productivity Distributions by Tenure for Workers Observed in Both Crops



Notes: We plot histograms of daily output (pieces) for workers observed working in both crops at some point in their tenure. 'Low Tenure' shows results for workers in their first 4 days on the farm. 'High Tenure' shows results for those with 10 or more days of tenure.

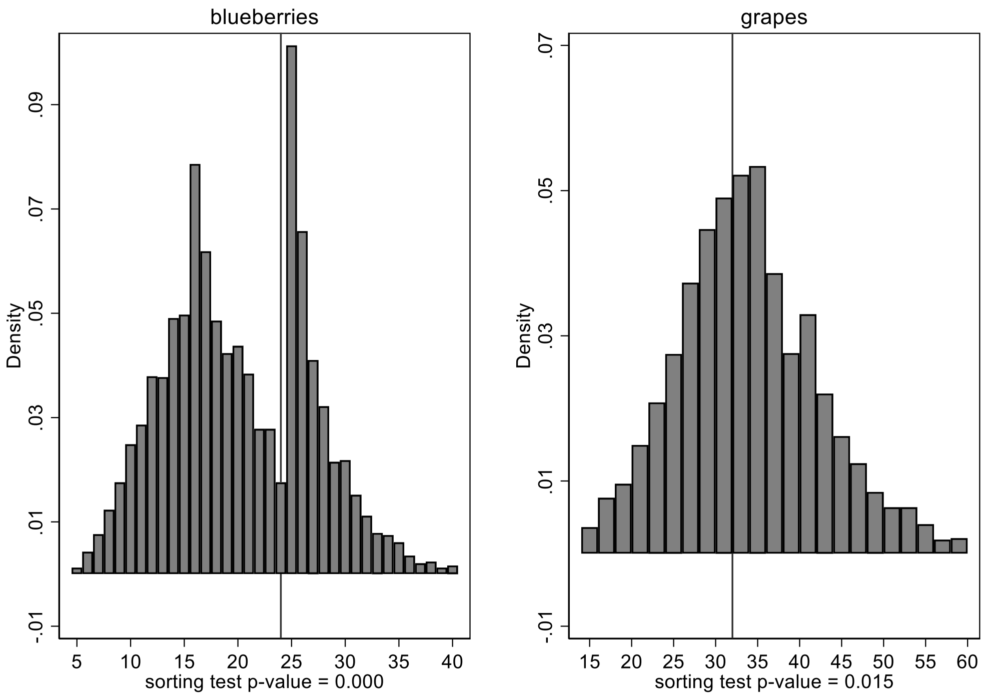
Appendix Table A1: Productivity Tenure Profiles by Starting Performance

	Starting Performance Quintile:				
	First	Second	Third	Fourth	Fifth
Blueberries	-0.197**	-0.396***	-0.425***	-0.536***	-0.769***
	(0.081)	(0.099)	(0.086)	(0.108)	(0.117)
Tenure/10	1.137***	0.381***	-0.298*	-0.141	-0.159
	(0.140)	(0.123)	(0.152)	(0.100)	(0.119)
$Tenure^2/1000$	-2.725***	-0.759***	0.792***	0.256	0.236
	(0.421)	(0.240)	(0.301)	(0.211)	(0.191)
Blueberries*Tenure/10	0.056	0.256*	0.465***	0.321**	0.302
	(0.132)	(0.142)	(0.155)	(0.156)	(0.182)
Blueberries*$Tenure^2/1000$	1.092**	0.050	-0.482	-0.497	-0.104
	(0.425)	(0.335)	(0.448)	(0.305)	(0.451)
Constant	-1.299***	-0.581***	-0.183*	0.297***	1.364***
	(0.127)	(0.116)	(0.100)	(0.082)	(0.104)
Observations	2,854	2,928	2,575	2,756	2,695
R-squared	0.458	0.304	0.186	0.208	0.219

Notes: *** p<0.01, ** p<0.05, * p<0.1; standard errors are clustered by worker. This table reports regressions of standardized daily performance (mean 0, standard deviation 1 within crop) for each quintile of mean 1-3 day performance. Regressions include controls for gender and date (see table 2). The sample is restricted to workers who are in their first crop and eventually survive for at least 10 days.

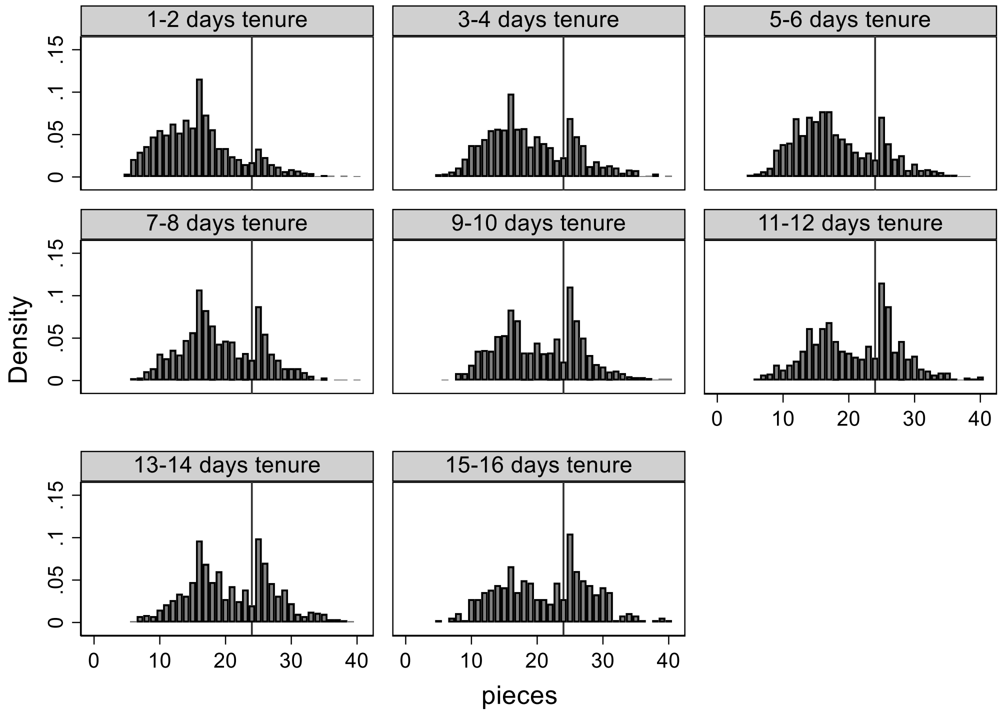
Online Appendix

Online Appendix Figure A1. Productivity distribution by crop, all workers (see Figure 4)



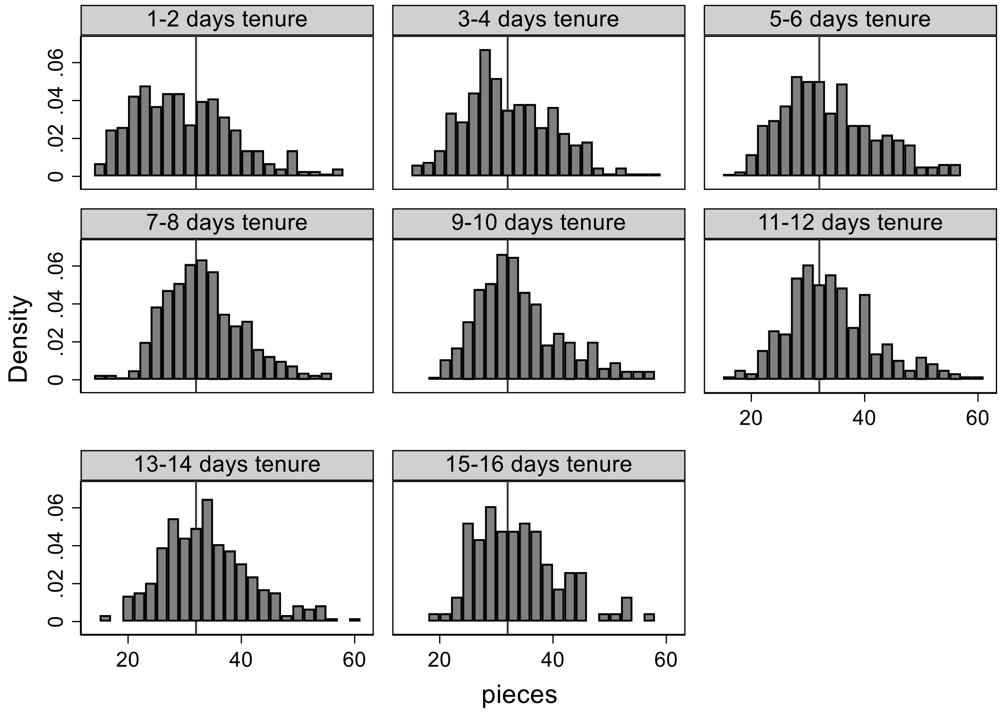
Notes: We plot histograms of daily output (pieces), separately by crop, for all workers regardless of completed tenure. The vertical lines show the threshold at which workers cross from the minimum wage to the piece rate regime. We report p-values for the Fransden (2017) heaping test, testing whether the performance distribution is smooth across the piece rate threshold.

Online Appendix Figure A2. Blueberry productivity by tenure, all workers (see Figure 5)



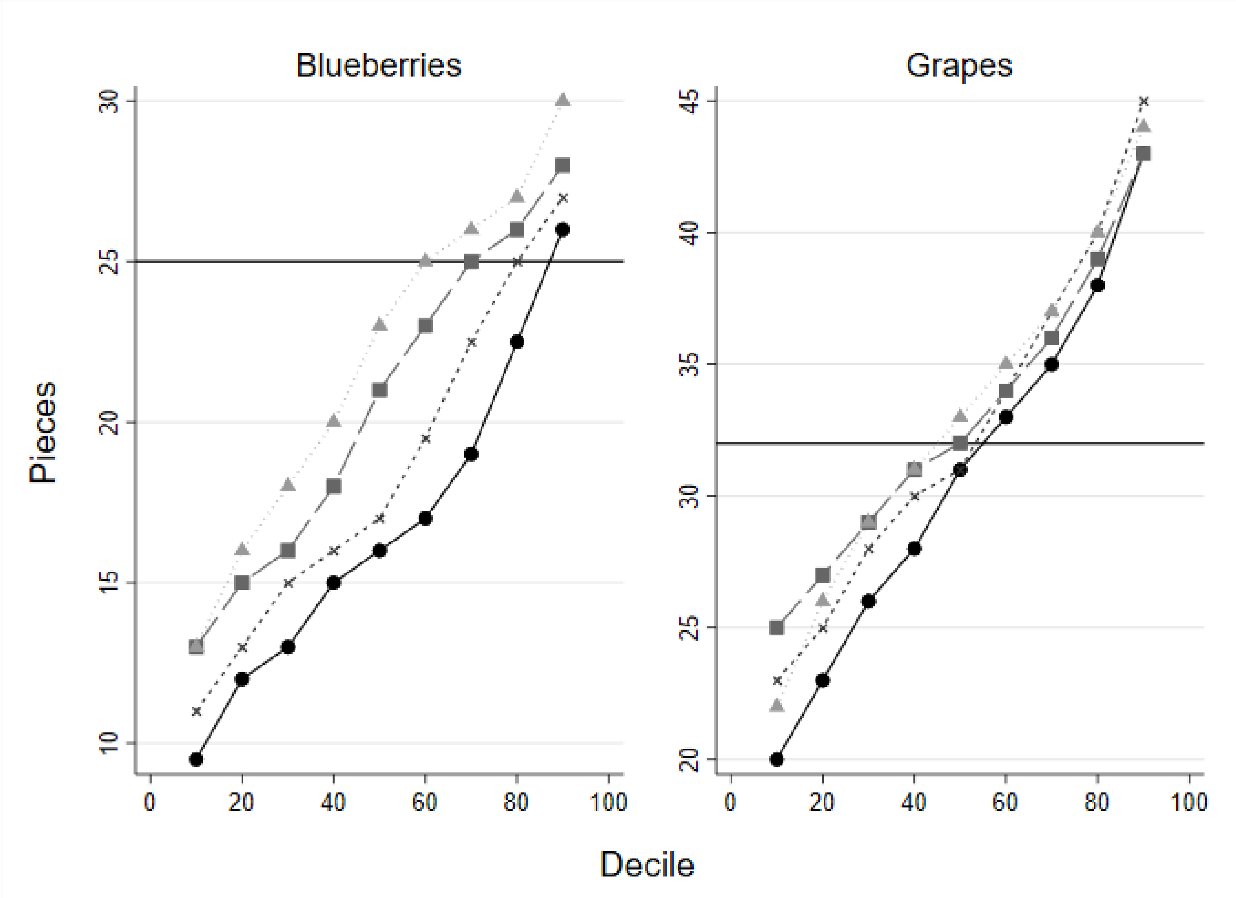
Notes: We plot histograms of daily output (pieces), separately by day of worker tenure. Graphs include all workers, regardless of completed tenure.

Online Appendix Figure A3. Grape productivity by tenure, all workers (see Figure 6)



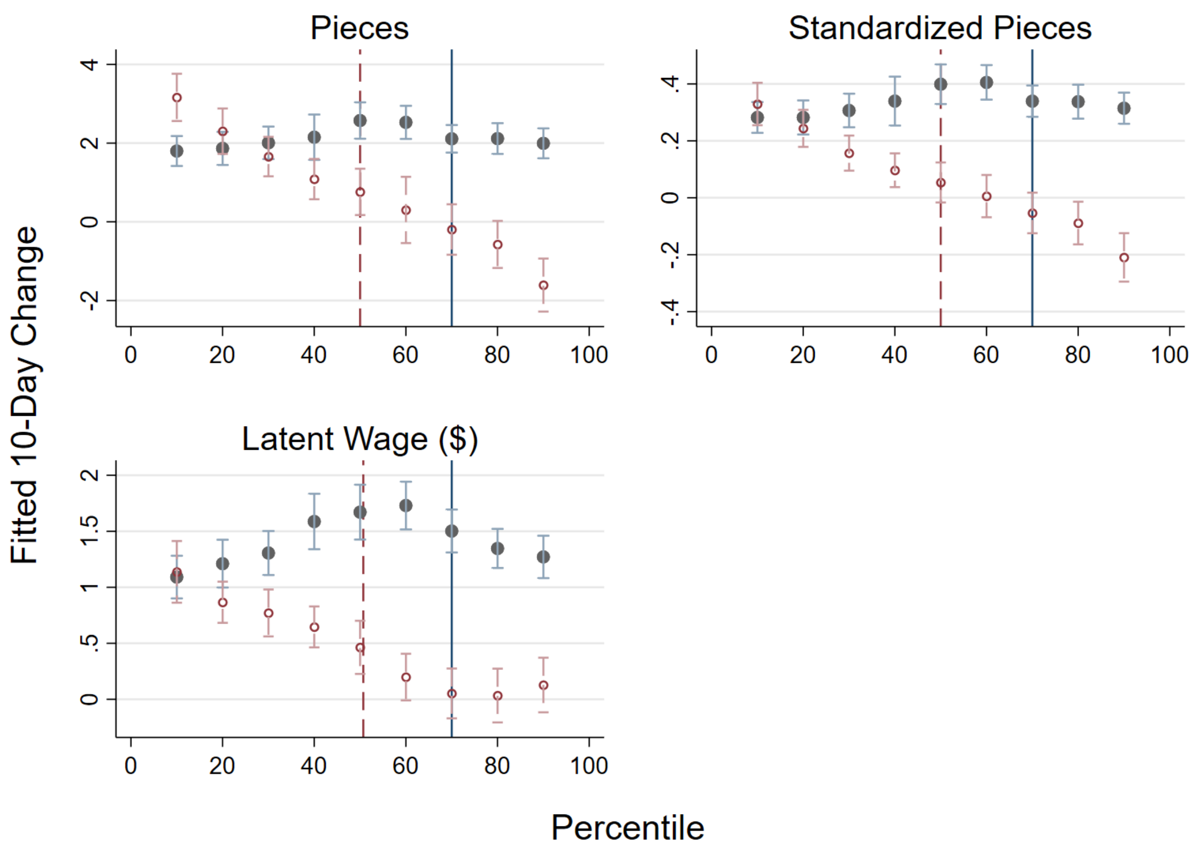
Notes: We plot histograms of daily output (pieces), separately by day of worker tenure. Graphs include all workers, regardless of completed tenure.

Online Appendix Figure A4: Performance Decile by Tenure and Crop, All workers (see figure 8)



Notes: Darker colors indicate newer workers. Circle = tenure < 4 days, X = tenure 4-6 days, square = tenure 7-9 days, and triangle = tenure 10 or more days. Figure plots the daily output associated with each decile in the distribution of pickers by crop and tenure group. There is no sample restriction based on completed tenure, though each line is restricted to the indicated tenure group, and we still exclude workers not in their first crop. The piece rate threshold for blueberry (grape) pickers is 25 (32), indicated with a horizontal line.

Online Appendix Figure A5: Fitted 10-Day Performance Growth, by Decile, Tenure and Crop, all workers (see Figure 9)



Notes: Solid = blueberries, hollow = grapes. Capped vertical lines indicate 90% confidence intervals. Using quantile regression, we estimate quadratic tenure profiles interacted with crop, for each performance decile, controlling for gender, day-of-week, week-of-harvest, and first/last harvest day. We plot fitted effects for the 10th day along with 90% confidence intervals. There is no sample restriction based on completed tenure, though we still exclude workers not in their first crop. Vertical lines indicate the point in the distribution where the piece rate thresholds are for each crop (blueberry line is solid, grape line is dashed). Standardized pieces (top right) are pieces normed to be mean 0, standard deviation 1, within crop. Latent Wage (bottom left) are the number of pieces times the piece rate ($0.50 for blueberries, $0.30 or $0.35 for grapes).

Online Appendix Table 1. Summary Statistics, all workers (see Table 1)

	Blueberries	Grapes
Pieces	20.13	32.99
	(6.73)	(8.81)
Daily Earnings ($)	68.76	65.24
	(6.05)	(1.80)
Hits piece rate threshold	0.34	0.56
	(0.47)	(0.50)
Tenure (days)	11.21	14.56
	(9.52)	(12.29)
Completed Tenure	22.28	28.85
	(13.44)	(15.71)
Survives ≥ 10 Days	0.85	0.94
	(0.36)	(0.24)
# worker-day obs	12,202	5,021
# unique days	57	44
# unique workers	984	367

Notes: Table reports means (and standard deviations) for the full sample of workers, regardless of completed tenure. Pieces is the number of blueberry buckets or grape boxes (as indicated) picked in the day. The piece rate threshold is 32 for blueberries and 25 for grapes. Daily earnings normalizes to an 8-hour work day.

Online Appendix Table A2. Probability of exceeding piece rate threshold, all workers (see Table 2)

Dependent Variable: Hit Piece Rate Threshold				
	(1)	(2)	(3)	(4)
Blueberries	-0.295***	-0.358***	-0.361***	-0.319***
	(0.026)	(0.029)	(0.029)	(0.029)
Tenure/10	0.046***	0.052*	0.052*	0.062**
	(0.010)	(0.027)	(0.026)	(0.031)
$Tenure^2/1000$		-0.013	-0.012	-0.003
		(0.053)	(0.052)	(0.060)
Blueberries*Tenure/10	0.099***	0.228***	0.229***	0.078**
	(0.015)	(0.033)	(0.033)	(0.036)
Blueberries*$Tenure^2/1000$		-0.379***	-0.379***	-0.163**
		(0.079)	(0.078)	(0.083)
Constant	0.482***	0.479***	0.490***	0.441***
	(0.022)	(0.026)	(0.031)	(0.032)
Observations	16,101	16,101	16,101	16,101
R-squared	0.088	0.095	0.095	0.179
Gender Control			X	X
Date Controls				X

Notes: *** p<0.01, ** p<0.05, * p<0.1. See table 2. We regress an indicator for whether the worker hit his piece rate threshold on the given day. The threshold for grapes is 32 pieces and the threshold for blueberries is 25. There is no sample restriction based on completed tenure, though we still exclude workers not in their first crop. Gender controls are indicators for male and female (with missing gender as the omitted category). Date controls include day of week, week of harvest, and first and last day of harvest indicators. Standard errors are clustered by worker.

Online Appendix Table A3: Productivity Tenure Profiles throughout the Productivity Distribution, All Workers (see Table 3)

Dependent Variable: Daily Output Quintile for Standardized Pieces				
Quintile:	20	40	60	80
Blueberries	-0.071	-0.263***	-0.306***	-0.465***
	(0.076)	(0.081)	(0.088)	(0.106)
Tenure/10	0.330***	0.139**	0.006	-0.164
	(0.072)	(0.069)	(0.096)	(0.113)
$Tenure^2/1000$	-0.384**	0.026	0.281	0.550**
	(0.160)	(0.132)	(0.208)	(0.268)
Blueberries*Tenure/10	-0.053	0.196*	0.423***	0.506***
	(0.088)	(0.106)	(0.102)	(0.116)
Blueberries*$Tenure^2/1000$	0.138	-0.298	-0.985***	-1.084***
	(0.207)	(0.276)	(0.228)	(0.279)
Constant	-1.202***	-0.571***	-0.228**	0.487***
	(0.083)	(0.088)	(0.093)	(0.111)
Observations	16,101	16,101	16,101	16,101
R-squared	0.153	0.182	0.175	0.169

Notes: *** p<0.01, ** p<0.05, * p<0.1. See table 3. We estimate quantile regressions of standardized daily performance (mean 0, standard deviation 1 within crop) for each indicated quantile. There is no sample restriction based on completed tenure, though we still exclude workers not in their first crop. Regressions include controls for gender and date (see online appendix table 1). Standard errors are clustered by worker.

Online Appendix Table A4: Productivity Tenure Profiles by Starting Performance, All Workers (see main appendix table A1)

Dependent Variable: Daily Output Quintile for Standardized Pieces					
	Starting Performance Quintile:				
	First	Second	Third	Fourth	Fifth
Blueberries	-0.117	-0.311***	-0.432***	-0.502***	-0.636***
	(0.071)	(0.088)	(0.072)	(0.099)	(0.102)
Tenure/10	1.220***	0.532***	-0.211**	-0.069	-0.149
	(0.136)	(0.140)	(0.095)	(0.101)	(0.115)
$Tenure^2/1000$	-2.954***	-1.087***	0.550***	0.159	0.240
	(0.411)	(0.284)	(0.201)	(0.211)	(0.187)
Blue*Tenure/10	-0.029	0.151	0.519***	0.378***	0.187
	(0.132)	(0.152)	(0.126)	(0.145)	(0.175)
$Blue*Tenure^2/1000$	1.330***	0.290	-0.549	-0.615**	0.045
	(0.428)	(0.363)	(0.407)	(0.291)	(0.462)
Constant	-1.375***	-0.652***	-0.154*	0.230***	1.242***
	(0.114)	(0.114)	(0.079)	(0.080)	(0.095)
Observations	3,319	3,098	3,195	3,178	3,183
R-squared	0.473	0.311	0.189	0.186	0.190
Full Controls	X	X	X	X	X

Notes: *** p<0.01, ** p<0.05, * p<0.1; standard errors are clustered by worker. This table reports regressions of standardized daily performance (mean 0, standard deviation 1 within crop) for each quintile of mean 1-3 day performance. Regressions include controls for gender and date (see table 2). The sample is no longer restricted to a completed tenure threshold, though we still restrict to workers who are in their first crop.