

Kahn, Lisa B.

Working Paper

Asymmetric information between employers

IZA Discussion Papers, No. 7147

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Kahn, Lisa B. (2013) : Asymmetric information between employers, IZA Discussion Papers, No. 7147, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/69339>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 7147

Asymmetric Information between Employers

Lisa B. Kahn

January 2013

Asymmetric Information between Employers

Lisa B. Kahn

*Yale School of Management
and IZA*

Discussion Paper No. 7147
January 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Asymmetric Information between Employers^{*}

Employer learning about workers' abilities plays a key role in determining how workers sort into jobs and are compensated. This study explores whether learning is symmetric or asymmetric, i.e., whether potential employers have the same information about worker ability as the incumbent firm. I develop a model of asymmetric learning that nests the symmetric learning case and allows the degree of asymmetry to vary, yielding testable implications for the prevalence of asymmetric learning. I then show how predictions in the model can be tested using compensation data. Using the NLSY, I test the model and find strong support for asymmetric information. I first exploit the fact that groups of workers differ in their variances in ability – based on economic conditions at time of entry into a firm – to show that incumbent wages track differences in ability distributions more closely than do outside firm wages. Second, I show that learning about ability is more symmetric for occupations that require more communication outside the firm. Finally, I show how to uncover the key parameter of interest in my model representing the degree to which information is asymmetric. My estimates imply that in one period, outside firms reduce the average expectation error over worker ability by roughly a third of the reduction made by incumbent firms. Thus outside firms retain sizeable expectation errors due to asymmetric information.

JEL Classification: D21, D83, J33

Keywords: employer learning, asymmetric information, personnel economics

Corresponding author:

Lisa Kahn
Yale School of Management
135 Prospect St
PO Box 208200
New Haven, CT 06510
USA
E-mail: lisa.kahn@yale.edu

^{*} I am grateful to Jason Abaluck, John Beshears, Richard Freeman, John Friedman, Carola Frydman, Bob Gibbons, Tal Gross, Kevin Hallock, Naomi Hausman, Mitchell Hoffman, Richard Holden, Guido Imbens, Radha Iyengar, Karthik Kalyanaraman, Fabian Lange, Steve Levitt, Barry Nalebuff, Emily Oster, Josh Pinkston, Giacomo Ponzetto, Uta Schönberg, Josh Schwartzstein, Mike Waldman, seminar participants at Harvard, IZA, University of Michigan, University of Minnesota, University of Pennsylvania, UC Santa Barbara, Vanderbilt, Yale, the 2008 SoLE meetings, the 2008 NBER Summer Institute, and especially Larry Katz, Claudia Goldin, Caroline Hoxby, and Thomas Lemieux for helpful comments and guidance.

1 Introduction

The sorting of workers into firms is an important aspect of the labor market. A significant part of this sorting involves both worker and firm learning about the quality of workers over time. For example, many patterns observed in the job mobility literature – that young workers change jobs often (Topel and Ward 1992) yet long-term employment relationships are common and the probability of remaining at a firm rises with tenure (Farber 1999) – are consistent with learning over time about overall or match-specific worker quality. However, the extent to which information about workers spreads across the market is less clear. Asymmetric information exists when incumbent employers have more information about worker quality than potential employers do. Consider a team of research assistants working for a professor. The professor can see exactly which worker performed each task and who contributed important ideas, yet the outside world observes only what appears on the bottom of an academic paper: the professor is grateful for the excellent research assistance of...

Under asymmetric information, inefficiencies can arise in both allocating workers to jobs and investing in worker human capital. First, the asymmetric information model implies that lower-ability workers are more likely to leave the firm (Greenwald 1986 and Gibbons and Katz 1991) since the incumbent employer but not the outside employer has learned that these workers are low ability. Higher quality workers might leave firms at inefficiently low rates. Second, Waldman's 1984 model assumes that outside firms can observe promotions but not wages and implies that a worker's outside option changes sharply upon promotion.¹ The incumbent firm must offer a large wage increase to keep a promoted worker so there are fewer promotions than would be optimal.² Third, several theoretical papers point out that under asymmet-

¹Other theoretical papers focusing on the promotion-as-signal hypothesis include Bernhardt (1995) which contains an extended version of the Waldman (1984) model with more empirical predictions and Zabochnik and Bernhardt (2001) which brings tournament theory and general human capital investment into the model.

²Similarly, Milgrom and Oster (1987) hypothesize that skills of minority workers are difficult for outside firms to discern, but promotions increase observability. Firms thus

ric information, workers could underinvest in general skills.³ When outside firms cannot observe workers' investments, the incumbent firm need not fully compensate the worker.⁴ It is necessary to understand the importance of asymmetric information to know whether these inefficiencies should be of concern.

This paper seeks to identify whether asymmetric information is prevalent in the labor market. That an employer learns more quickly than the outside market about its own workers' abilities is intuitive, yet little empirical evidence exists. A small number of papers provide empirical support for asymmetric learning but usually require strong assumptions or look in specialized settings.⁵ Several papers use AFQT score to provide evidence that employer learning is indeed prevalent, but they assume that learning is symmetric across employers.⁶ A few papers extend this methodology to the asymmetric learning case. For example, Schönberg (2007), the closest to the current paper, looks at whether the relationship between AFQT and earnings increases with firm tenure, controlling for the increased effect of AFQT across levels of experience. This would imply that incumbent employers learn more about worker productivity than does the rest of the market; however, with noisy estimates, she does

underpromote minority workers to retain their informational advantage.

³See, for example, Chang and Wang (1996), Katz and Ziderman (1990) and Waldman (1990).

⁴Alternatively, Acemoglu and Pischke (1998) show that providing general training might be optimal under asymmetric information because it allows high-quality workers to signal their type by selecting into a firm that provides such training.

⁵Gibbons and Katz (1991) show that laid off workers experience larger earnings losses, relative to workers who lost their jobs due to plant closings, consistent with the presence of asymmetric information and a layoff being a negative signal of ability. DeVaro and Waldman (2012) test an extended version of the Waldman (1984) model, using proprietary data from a single firm, and find support for asymmetric information. Using data on apprentices in Germany, Acemoglu and Pischke (1998) link employer-provided training to asymmetric information and von Wachter and Bender (2006) show that workers released by firms are negatively selected.

⁶In their canonical paper, Farber and Gibbons (1996) argue that the Armed Forces Qualifying Test (AFQT), administered to respondents in the NLSY in 1980, is unobservable to employers yet highly correlated with ability. They show that the effect of AFQT on wages becomes stronger over time and that of education becomes weaker, providing evidence that employers learn about worker ability. Altonji and Pierret (2001) expand this methodology to study statistical discrimination and Lange (2007) provides an estimate for the speed of employer learning.

not find a significant AFQT-tenure-earnings profile and therefore fails to reject the symmetric learning model, except possibly among college graduates.⁷ Though the AFQT approach presents an obvious advancement in the study of employer learning, it also brings with it an undesirable set of assumptions: first that the econometrician has information the firm does not but would have been valuable to collect; second, since AFQT was measured prior to labor market entry, this approach cannot distinguish between employer learning and models of idiosyncratic human capital accumulation.⁸ Furthermore, the papers using AFQT to test for asymmetric learning identify off of the interaction of AFQT score and tenure, thus additionally requiring strong assumptions about worker mobility.

In this paper, I contribute a new methodology for identifying employer learning, using compensation. My approach arguably lays fewer demands on a dataset and does not rely on the assumptions required by the AFQT approach. I derive a learning model with endogenous mobility and asymmetric information that nests symmetric learning as a special case. I then test the model using the National Longitudinal Survey of Youth 1979 (NLSY) and find strong support, both economically and statistically, in favor of asymmetric learning. Finally, I uncover the key parameter of interest in my model, the degree to which employer learning is asymmetric, using a linear regression framework. This is the first such estimate, to my knowledge. I conclude that asymmetric information between employers is prevalent in the labor market and that the effects are large in magnitude.

Specifically, I consider a two-period model in which workers have stochastic ability that is initially unobservable to all parties. Workers randomly match to firms at the beginning of period 1 and information about their ability is

⁷Zhang (2007) and Pinkston (2009) both employ similar methodologies, comparing the returns to AFQT across tenure relative to experience. Zhang provides additional empirical tests using histories of job changes, while Pinkston models raids and bidding wars, both finding evidence in favor of asymmetric information.

⁸Kahn and Lange (2012) use personnel data containing a panel of pay and performance measures to study employer learning and productivity evolution. They find evidence in favor of employer learning but also find that heterogeneous human capital accumulation has a sizeable impact on wages.

revealed. This model nests both the symmetric and asymmetric learning cases by allowing for outside employers to receive a noisier signal of worker ability than that of the incumbent employer. Based on new information, profit maximizing firms make wage offers. Workers then choose whether to stay at the incumbent firm or take an outside wage offer, based on the size of the offers and the realization of a random preference shock specific to the incumbent firm (a la Acemoglu and Pischke 1998). In the second period workers earn either the incumbent firm's wage offer or the outside offer and then retire.⁹

This model yields two important implications about the pay change distributions of movers and stayers. These are of interest since wage changes reflect how much an employer has learned between periods and a pay change distribution reflects learning about a group of workers. First, the variance of wage changes is increasing in the variance in ability, but under asymmetric information this relationship is stronger for wage changes of stayers than for movers – this reflects the fact that incumbents have learned more about differences in worker ability between periods than have outside employers. Second, the variance of wage changes is declining in the degree of asymmetry in the market, but this relationship is stronger for wage changes of movers than for stayers – the greater the degree of asymmetry, the less outside firms have learned about worker ability relative to the incumbent, and therefore the more compressed their wage offers will be.

I test the model using the National Longitudinal Survey of Youth 1979. These tests require variation in the dispersion of worker ability and the degree of asymmetry. For cross-group differences in the variance in worker ability I use labor market conditions at time of entry into a firm. I provide evidence that workers who enter firms in recessions have a *lower* variance in ability than workers who enter firms in booms. During a recession, voluntary quits fall substantially;¹⁰ hence, the share of job-seekers who are unemployed or leaving

⁹My modeling approach draws on Schönberg (2007), although her model has only two types of workers and a binary signal to outside employers, while mine has a continuum of types and a normally distributed signal.

¹⁰For example, during the Great Recession, the quit rate fell by nearly 50% (BLS, Job Openings and Labor Turnover).

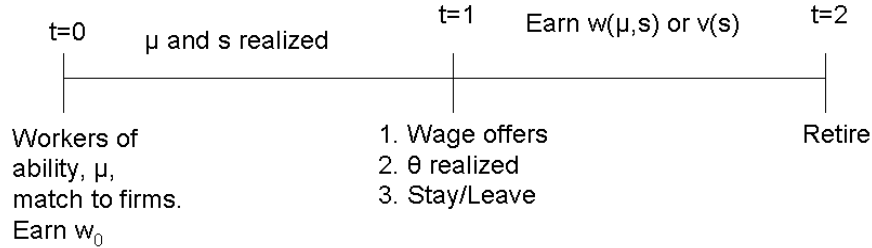
bad jobs increases relative to boom-times. In consequence, the variance of worker ability will be lower for workers who enter firms in recessions – these workers are more likely to be of lower quality. Consistent with their lower variance in ability, workers who enter firms in recessions do indeed have a lower variance in wage changes within the firm but less so when they subsequently switch firms. This finding supports the asymmetric learning model because it indicates that outside wage offers are less linked to worker ability.

As a source of variation in the degree of asymmetry, I compare workers across job types that vary in the importance of communicating to people outside the firm, a classification created using O*NET task measures. I hypothesize that information will spread more symmetrically across employers when jobs involve more external interaction. I find that, consistent with my asymmetric information model, the variance in wage changes is smaller in the more asymmetric markets (with less outside communication) and this relationship is smaller in magnitude for stayers.

Finally, I use my estimates and some plausible calibration assumptions to quantify how relevant asymmetric information is in the economy as a whole. I find signals of worker quality received by outside firms are substantially noisier than what incumbents observe. I use this to calculate the differential speed of employer learning (as in Lange 2007) by outside firms, relative to the incumbent. I find that in one period, outside firms can reduce their initial expectation errors over worker quality by roughly a third of what the incumbent firm can reduce theirs by. Thus, asymmetric information, has a sizeable impact on expectation errors of firms.

The remainder of the paper proceeds as follows. Section 2 presents a theoretical model and derives the predictions to be tested in the paper. Section 3 shows how to empirically implement these predictions. Section 4 describes the data and the empirical methodology. Section 5 presents the results and shows that the predictions of the asymmetric information model are supported in the data. Section 6 concludes.

Figure 1: Timeline



2 Theory

The purpose of this section is to motivate the two comparative statics tested in the paper. In order to keep the model tractable, I make several simplifying assumptions. In particular, there is no match quality or human capital accumulation and the worker's turnover decision is reduced to a function of just a few variables, described below. In the next section, I discuss how I can empirically estimate the parameters of interest, despite these simplifying assumptions.

2.1 Basic Setup and Timing

In this two-period model, I assume a competitive labor market where firms are profit maximizing and both workers and firms are risk neutral with no discounting. The timing of events is summarized in figure 1. At time 0, worker i draws ability μ_i from a distribution with cumulative distribution and probability density functions Φ and ϕ . To simplify the analysis, I assume that ability is normally distributed, i.e., $\phi(\cdot) \sim N(m, \sigma^2)$.¹¹ I also assume that output equals ability in all firms in all periods. Thus there is no match quality or human capital accumulation after labor market entry.

¹¹The normality assumption is not important. What is important is that the ability distribution not be too skewed. The analysis carries through for many common distributions with some skew, such as pareto, log normal and exponential.

Also at time 0, workers are randomly matched to firms. Call the hiring firm the incumbent. At that point ability is unobservable to all parties, both workers and firms. Thus all workers earn the same first period wage, w_0 .

During period 1, information is revealed. I assume the incumbent firm learns μ_i perfectly, while outside employers do not. Instead, they observe a public, noisy signal of ability for each worker, s_i .¹² Let $s_i = \mu_i + \epsilon_i$ where $\epsilon \sim N(0, r\sigma^2)$, $r \geq 0$ and $\epsilon \perp \mu$. If $r = 0$ then $s = \mu$ with certainty and we are in the perfectly symmetric learning case. As r approaches ∞ , the signal becomes meaningless and we approach the perfectly asymmetric learning case. From Baye's Rule, the posterior on ability, μ , conditional on the realized signal, s , $\sim N(\frac{mr+s}{1+r}, \frac{r\sigma^2}{1+r})$ (DeGroot 2004).

Throughout the paper, I will be primarily interested in $X = \frac{r}{1+r}$, which is strictly increasing in r and ranges between 0 and 1. X has a natural interpretation as the weight outside firms place on their prior when updating beliefs about worker ability after receiving a new signal, or 1 minus the speed of employer learning defined in Lange (2007). I call X the degree of asymmetry because it represents the relative noisiness of the signal outside firms receive.

Based on this information, firms make wage offers. The incumbent firm's strategy is a mapping $w(\mu, s) : \mathbb{R}^2 \rightarrow \mathbb{R}$ while an outside firm's strategy is a mapping $v(s) : \mathbb{R}^2 \rightarrow \mathbb{R}$, both yielding period 2 wage offers. I further assume that outside firms cannot observe incumbent wage offers, otherwise ability, μ , could be perfectly inferred.

Mobility is driven by wage offers and a random disutility shock, θ_i , which workers receive during period 1 (following Acemoglu and Pischke 1998). I assume θ is uniformly distributed over the interval $[-\bar{\theta}, \bar{\theta}]$ and is independent of μ and s . It is the worker's ex-post evaluation of the workplace (e.g., dislike of coworkers) and workers must bear this cost if they stay with the incumbent, or get a new draw (with mean 0) if they move firms. Workers maximize income minus the disutility shock in making quit decisions. θ is meant to represent the level of random turnover in the economy, otherwise moving would be perfectly

¹²Rather than a public signal, I could instead assume that outside firms receive a private signal but incumbent firms can react to their wage offers.

revelatory of the incumbent’s wage offer so no one would find it worthwhile to move. I assume that θ cannot be credibly revealed to employers and therefore cannot be incorporated into wage offers.

2.2 Calculating the Equilibrium

The perfect Bayesian equilibrium (PBE) consists of an incumbent firm wage schedule, $w^*(\mu, s)$, an outside firm wage schedule, $v^*(s)$, and the worker’s quit decision, such that: (1) the incumbent firm maximizes profits conditional on outside firm behavior, (2) outside firms maximize profits subject to beliefs about worker turnover behavior, (3) workers maximize income minus θ in making their quit decision and (4) all beliefs are consistent with these strategies.

In the PBE, incumbent firms take the outside wage offers as given and maximize expected profits for each (μ, s) pair, calculating the expectation over all realizations of θ . Competition dictates that outside firms set a wage schedule, $v(s)$, equal to a worker’s expected productivity conditional on leaving the firm and on s .¹³ Thus the solution for $v^*(s)$ falls straightforwardly from Baye’s Rule. Both wage offers are represented below and derived in Appendix A.1.

$$v^*(s) = mX + s(1 - X) - c \tag{1}$$

$$w^*(\mu, s) = \frac{1}{2}(\mu + v^*(s) - \bar{\theta}) \tag{2}$$

Equation 1 shows that $v^*(s)$ equals the posterior expectation of μ conditional on s (DeGroot 2004) – simply a weighted average of the prior and the new signal – minus a constant. The constant reflects that outside firms also update on the information revealed by a worker’s willingness to accept the outside wage offer. The value of the constant is unimportant here since we are concerned with variances in pay. However, it is worth noting that $c > 0$ (see appendix A.1). Outside wage offers, equaling expected productivity con-

¹³The wage offer will converge on expected productivity if outside firms undergo Bertrand competition, for example.

ditional on s and moving, are thus lower than the expectation of μ given s unconditional on moving ($mX + s(1 - X)$). Therefore movers are on average negatively selected, a point made clearer in the discussion of the incumbent firm's wage offer, next.

In equation 2 the incumbent firm's wage schedule is the average of true productivity and outside firm expectations (plus a constant related to the distribution dictating random turnover). This formulation of the incumbent wage offer is quite intuitive. Having learned μ , the incumbent holds an information rent over the worker, the value of which is $\mu - w^*(\mu, s)$. The firm could set the wage equal to the worker's outside option, $v^*(s)$, and capture the entire information rent. However it then risks losing workers who have a high realization of θ , the disutility shock. Instead the firm could set the wage equal to μ and have a higher chance of outweighing θ for high quality workers. The firm would thus be more likely to keep high quality workers, but would lose the entire information rent. Facing these trade-offs (wanting to keep high quality workers while still capturing some of the value of the information rent), the firm chooses to set pay equal to the average of the two.¹⁴

Therefore incumbent wage offers are linked to true ability, μ , even though outside firms cannot observe it. Furthermore, wage offers are increasing in μ , implying the worker's probability of staying is also increasing in μ . Thus this model yields that classic lemons effect (a la Greenwald 1986) that worse workers are more likely to leave.

Finally, the period 1 wage can be calculated by adding a zero ex-ante profit condition (see appendix A.2). Since at the beginning of period 1 ability is unknown, w_0 is simply a constant.

2.3 Comparative Statics

In this section, I derive comparative statics relating to the variance of wage changes. This is of interest because the wage change incorporates new information firms have learned about worker ability. The degree to which this

¹⁴The $\frac{1}{2}$ coefficient comes out of the uniformity assumption over θ . This implies that a worker's probability of staying is linear in the wage gap, $w^*(\mu, s) - v(s)$.

distribution is correlated with the true underlying ability distribution reflects how well firms have learned. Note equations 1 and 2 take as arguments μ and s . However, when calculating their variances, I integrate over the actual probabilities that a given (μ, s) pair accepts that offer. Therefore, the propositions derived below hold for the *accepted* wage distribution, not the distribution of wage offers.

The change in wages between periods 1 and 2 for workers who stay with the incumbent firm, Δw^{stay} , is by definition, $w^*(\mu, s) - w_0$, while the wage change for movers, Δw^{move} , is $v^*(s) - w_0$. In appendix A.3, I calculate the variance of Δw^{stay} among those who accept $w^*(\mu, s)$ and the variance of Δw^{move} among those who accept $v^*(s)$.

For ease of notation, I will express the variance of wage changes for stayers (equation 8 in the appendix) as V^{stay} and the variance of wage changes of movers as V^{move} (equation 9 in the appendix). A closed form solution to $V^{stay} - V^{move}$ is expressed in appendix A.3 (equation 10), though here I simply note that it is a function of σ^2 , X , and $\bar{\theta}$ (respectively, the variance in ability, the degree of asymmetry, and the distribution parameter for the random disutility shock). To see how the variance in wage changes is affected by the underlying variance in ability across movers and stayers, I need to sign expression 3.

$$\partial_\sigma = \frac{\partial(V^{stay} - V^{move})}{\partial\sigma^2} \quad (3)$$

In appendix A.3, I show that ∂_σ is always *positive*, i.e., the derivative is larger for the variance in wage changes of stayers than for movers. Under asymmetric information, the incumbent firm has learned more about worker ability so the distribution of wage changes can more closely track the distribution of ability. Intuition for this result can be gleaned from considering equations 1 and 2, respectively $v^*(s)$ and $w^*(\mu, s)$. Variation in the incumbent wage offer comes from μ and s , while variation in the outside wage offer comes only from s . Recall s is noisily related to μ , with the amount of noise being dictated by X , the degree of asymmetry. Therefore the variation in these wage offers will be positively related to variation in μ (i.e., σ^2). However, since the outside

wage offer is only noisily related to μ , its variance will be less related to σ^2 . Indeed, as X approaches 1 and we approach the pure asymmetric case, outside firms place no weight on the new signal, so the wage change for movers has no variance. At the same time, it is easy to show that when $X = 0$, all firms have the same information set, and outside firms place full weight on the new signal. Since, in this case, the new signal equals true ability, μ , perfectly, wage distributions of both incumbents and outside firms will be equally related to σ^2 . Proposition 1 summarizes this finding.

Proposition 1 *Under asymmetric information, the variance in wage changes responds more to differences in the variance of ability for stayers than for movers ($\partial_\sigma > 0$). Under symmetric information, the response is the same across movers and stayers ($\partial_\sigma = 0$).*

To see how the variance in wage changes is affected by changes in X , the degree of asymmetry in the market, across movers and stayers, we need to sign expression 4.

$$\partial_X = \frac{\partial(V^{stay} - V^{move})}{\partial X} \quad (4)$$

In appendix A.3, I show that ∂_X is always positive. A higher degree of asymmetry, X , compresses the distributions of wage changes for both movers and stayers, but more so for movers. The intuition is that the outside wage offer, $v^*(s)$, is less related to true ability when information is more asymmetric, that is when s is more noisily related to μ . Since $v^*(s)$ is a component of period 2 wages for both movers and stayers, period 2 wages will be less linked to true ability, implying wage changes will be less dispersed. However, this compression will have a larger impact on the variance for movers. This is because variation in wages of movers is entirely driven by variation in $v^*(s)$, which is only one component of the variance for stayers (recall from equation 2 that the period 2 wage for stayers is the average of μ and $v^*(s)$). This result is summarized in proposition 2.

Proposition 2 *The variance in wage changes for movers responds more negatively to the degree of asymmetry than does the variance for stayers ($\partial_X > 0$).*

3 Implementation

The goal in this section is to show how I can empirically identify expressions 3 and 4. This requires credible variation in both σ^2 and X , two variables not readily observable in most datasets. In addition, I need to discuss factors outside the scope of the model that could influence these expressions empirically.

Suppose we have two groups of workers, one with a higher variance ability distribution than the other. Let σ_H^2 and σ_L^2 express the variances in ability for each group, where $\sigma_H^2 > \sigma_L^2$. Similarly, suppose there are two types of jobs that vary in the degree to which outside firms can observe worker quality. Let X^A and X^S represent the degree of asymmetry for the two groups, where $X^A > X^S$, i.e., the noise term on the signal of worker quality that outside firms observe is higher variance for the former, relative to the latter. Finally, suppose each type of worker can be in either type of job.

Equation 5 expresses the variance in wage changes for both worker types ($i \in H, L$) and both job types ($j \in A, S$). It uses the indicator variable, I^{stay} (equaling one if the worker stayed at the firm between periods) to combine the expressions for the variances in pay changes for movers and stayers. Equation 5 further contains three additional terms: γ^{stay} , γ^{σ_i} , and γ^{r^j} . I explain these terms next.

$$Var(\Delta w) = I^{stay} * V^{stay} + (1 - I^{stay}) * V^{move} + \gamma^{stay} + \gamma^{\sigma_i} + \gamma^{X^j} \quad (5)$$

First, γ^{stay} allows the variance in pay changes to be different for stayers than for movers for reasons beyond the scope of the model. For example, changing firms also involves a new mix of amenities, a new realization of match quality and changing compensating differentials, all of which will impact pay and will likely cause the variance in pay changes to be larger for movers than for stayers. Also, search theory motivates an additional friction resulting in

inefficient worker mobility that would have an impact on pay. Finally, the asymmetric information model predicts that movers will be negatively selected. However, evidence suggests that workers often move to better opportunities and experience wage growth upon doing so (Topel and Ward 1992), or learn about their own abilities and sort into better matches over time (Gibbons, Katz, Lemieux and Parent 2005). All of these factors and more are incorporated into γ^{stay} . This is a sizeable advancement over the previous literature since, to the best of my knowledge, none of the previous papers relying on the interaction of AFQT and tenure to test asymmetric information account for these factors. My key assumption is the weaker one that these turnover factors do not interact with worker or job type.

In addition, γ^{σ_i} and γ^{X^j} allow the variances in pay changes to be different for workers with different variances in ability and jobs that differ in the degree to which information is asymmetric. One motivation for these fixed effects is that groups may accumulate human capital at differing rates, which would result in differing degrees of pay dispersion.¹⁵ Their importance will become clear below, when I discuss the empirical analogues for groups that differ in variance in ability and jobs that differ in the degree of asymmetry.

To test propositions 1 and 2 above, we need to empirically estimate expressions 3 and 4, which represent the differences in the derivatives of pay changes between movers and stayers with respect to σ^2 and X , respectively. Let us begin with the identification of ∂_σ (equation 3, above).

Equations 6-9 show the formulas for the variances in pay changes for four groups: those in the high variance ability distribution who stay, those in the high variance ability distribution who move, those in the low variance ability distribution who stay and those in the low variance ability distribution who move, respectively. These come directly from equation 5. Importantly, as

¹⁵For example, if human capital investments and ability are complements so that more able people should obtain higher levels of human capital (a la Gibbons and Waldman 1999 or Acemoglu and Pischke 1998), then we would expect groups of workers with different ability distributions (σ_H^2 and σ_L^2) to invest at different rates. Furthermore, as discussed above, when information is asymmetric, workers have less incentive to invest in general human capital. We would therefore expect different investment levels across groups with differing degrees of asymmetry (r^A and r^S).

indicated, I assume X and $\bar{\theta}$ are held constant and γ^{stay} does not vary across high- and low-variance ability groups.

$$Var(\Delta w)_{stay}^H = V^{stay}(\sigma_H^2, X, \bar{\theta}) + \gamma^{stay} + \gamma^{\sigma_H} + \gamma^X \quad (6)$$

$$Var(\Delta w)_{move}^H = V^{move}(\sigma_H^2, X, \bar{\theta}) + \gamma^{\sigma_H} + \gamma^X \quad (7)$$

$$Var(\Delta w)_{stay}^L = V^{stay}(\sigma_L^2, X, \bar{\theta}) + \gamma^{stay} + \gamma^{\sigma_L} + \gamma^X \quad (8)$$

$$Var(\Delta w)_{move}^L = V^{move}(\sigma_L^2, X, \bar{\theta}) + \gamma^{\sigma_L} + \gamma^X \quad (9)$$

It should already be clear that we will be able to difference out most of these terms. The difference-in-differences expression is shown below. To get to equation 10, I approximate the change in $V^{stay} - V^{move}$ associated with an increase in σ^2 by the derivative of $V^{stay} - V^{move}$ with respect to σ^2 times the size of the increase, where here $\Delta\sigma^2 = (\sigma_L^2 - \sigma_H^2)$.

$$\begin{aligned} dd_\sigma &= (Var(\Delta w)_{stay}^L - Var(\Delta w)_{move}^L) - (Var(\Delta w)_{stay}^H - Var(\Delta w)_{move}^H) \\ dd_\sigma &\approx \partial_\sigma * \Delta\sigma^2 \end{aligned} \quad (10)$$

Since $\Delta\sigma^2 < 0$ and proposition 1 tells us $\partial_\sigma > 0$ under the asymmetric learning model, we should expect an empirical estimate of $dd_\sigma < 0$. That is, the decrease in the variance in wage changes seen in the low variance group compared to the high variance group is larger in magnitude for stayers than for movers. In contrast, under symmetric information, period 2 wage offers are equally linked to ability for both movers and stayers, hence so are the variances and $dd_\sigma = 0$.

Now let's turn to identifying ∂_X (expression 4, above). We can write a set of equations, analogous to 6-9, where we vary X (between X^A and X^S) and hold constant σ^2 . I can then derive the difference-in-differences estimate in a similar manner. This is shown below, where equation 11 approximates the change in $V^{stay} - V^{move}$ associated with an increase in X by the derivative with respect to X times the size of the increase, where $\Delta X = (X^S - X^A)$. Since $\Delta X < 0$ and proposition 2 tells us $\partial_X > 0$ under the asymmetric learning

model, we should expect an empirical estimate of $dd_X < 0$.

$$\begin{aligned} dd_X &= (Var(\Delta w)_{stay}^S - Var(\Delta w)_{move}^S) - (Var(\Delta w)_{stay}^A - Var(\Delta w)_{move}^A) \\ dd_X &\approx \partial_X * \Delta X \end{aligned} \tag{11}$$

Having shown how to identify the expressions of interest using a difference-in-differences strategy, I now describe the data and show how to empirically estimate dd_σ , dd_X , and how I use these estimates to quantify X .

4 Data and Methodology

The primary dataset used in this paper is the National Longitudinal Survey of Youth (NLSY) from 1979-2000.¹⁶ These data allow me to follow a nationally representative sample of workers over a long period of time. The sample is restricted to the male cross-section sample (3,003 respondents), omitting women and the oversamples.¹⁷ The NLSY is particularly useful because it allows respondents to list up to five employers in the past year and these employer-employee relationships can be tracked across survey years. In each year, the survey records wage, occupation, industry, tenure for each job, and reason for having left the job if the worker is no longer there. For a job spell to be included in the sample, all wage observations at that employer must reflect full-time, non-enrolled, non-self-employed work of workers who have completed their transition from school into the labor force.¹⁸ I allow workers to hold only one such job at a time, excluding the shorter duration job for

¹⁶Starting in 1979, 12,686 youths between the ages of 14 and 21 were interviewed annually until 1994, then biannually thereafter. The survey contained a nationally representative sample, as well as oversamples of blacks, Hispanics, the military and poor whites.

¹⁷Women are excluded because it is important to isolate workers with strong attachment to the labor market.

¹⁸Transition is defined similarly to Farber and Gibbons (1996) and Schönberg (2007) as primarily working (working full time for at least half the year) for two consecutive survey years. At the beginning of this two-year spell, the worker is assumed to be post-transition. Roughly 93% of the male cross-section sample has had a full-time, non-enrolled, non-self-employed job in the sample period. Of those roughly a quarter of people never transition to the labor market and are thus dropped from the sample.

those held concurrently. The data also include the Armed Forces Qualifying Test (AFQT) score, an aptitude test administered to all NLSY respondents in 1981.¹⁹ This test has been shown to correlate with earnings, but is likely unobservable to employers, making it a useful variable in past studies of employer learning.²⁰ Though I use the AFQT as supporting evidence, my paper provides a new methodology that does not rely on the AFQT and the several assumptions required of that approach. Appendix table B.1 summarizes the creation of the main variables used in this analysis.

Differences in σ^2 In order to test proposition 1, that the variance in wage changes is more closely linked to the variance in ability for stayers than for movers, it is necessary to identify groups with different variances in ability. For this, I use labor market conditions at time of entry into a firm. Specifically, I use the national unemployment rate in the year the worker entered the firm, divided into quartiles over the sample period.²¹ I exploit the empirical phenomenon that workers who enter firms during worse economies have a lower variance in observable characteristics (and therefore, plausibly, lower variances in unobserved ability) than workers who enter firms in booms. This is due to the fact that job seekers are negatively selected in recessions relative to boomtimes. This negative selection also implies a lower variance, for most ability distributions.

To build intuition, consider the following broad categorization of job seekers: 1. voluntary movers, 2. involuntary movers and the unemployed and 3. new labor market entrants. The first group are positively selected, the second group are negatively selected and the third group is probably close to random.²² During a recession, voluntary quits fall dramatically, while invol-

¹⁹Since respondents took the test at different ages, I use an age-adjusted score.

²⁰See Lange (2007) for a nice discussion of the unobservability of AFQT to employers. In addition Farber and Gibbons, Altonji and Pierret, Schönberg, and Pinkston among others, use AFQT to test employer learning models.

²¹The cutoffs for the quartiles are the following: <5.5, 5.5-6.1 6.2-7.2, >7.2. I used the unweighted set of national unemployment rates from 1970-2000 to calculate the cutoff points. This is preferable to using the sample-weighted population of job-entry cohorts because of the different patterns of job mobility over the business cycle.

²²Kahn (2010) finds that workers who graduate from college in bad economies earn sub-

untary separations rise. For example, in the Great Recession, the quit rate fell by nearly 50%, from a peak of 2.3 in December 2006, to a trough of 1.2 in September, 2009, while the layoff and discharge rate rose from 1.2 to 2.0 over a similar period.²³ Thus the composition of job seekers is more strongly loaded on the second and third groups. Negative selection in the second group may become less pronounced, since in order to make larger workforce reductions, firms may have to layoff workers who are higher up on the ability ladder. However, empirically, it looks as though the reduction in voluntary quits is the stronger force.

Figure 2 helps to illustrate this point for the top and bottom national unemployment rate quartiles, which should reflect the starkest comparison. Here I show kernel densities of age-adjusted AFQT score for each group, as well as vertical lines at their means. As can be seen the high unemployment rate group (solid red line) has both a lower mean value and a more compressed distribution.

Table 1 shows that these differences are statistically significant. Here I show means and variances for a range of characteristics across job-entry unemployment rate quartiles. The table includes highest grade completed, age-adjusted AFQT score and a log wage residual from the first wage observation at the employer (though very similar findings are obtained from any number of other variables contained in the NLSY).²⁴ These are useful measures in that education is observable to both the employer and the econometrician, while AFQT score is likely only observable to the econometrician and a first wage residual incorporates information available to the employer but not the econometrician. Patterns found in these measures, which are to varying degrees observable, are suggestive that similar patterns would hold with unobserved ability as well.

stantially lower wages throughout their careers. She shows that these effects are not driven by selection into graduation years. Therefore it is likely that the group of new labor market entrants does not systematically differ in quality across booms and busts.

²³These statistics come from the Bureau of Labor Statistics' Job Openings and Labor Turnover Survey (JOLTS). Rates are seasonally adjusted and reflect the number of monthly exits as a percent of total employment.

²⁴Log wage is residualized on a quadratic in age, marital status, and education, geographic region, urban status, black and Hispanic dummies.

Panel A shows differences in means, while panel B shows differences in variances, where the first row in each panel compares the highest quartile unemployment rate (worst economy) to the lowest, yielding the sharpest contrast.²⁵ As can be seen, workers entering firms in economic slumps have less education, lower test scores and a lower first wage residual. They also have lower variances along all of these dimensions, and these differences are all statistically significant at the 10% level or better. In the next section, I use these job-entry cohorts as a source of variation in σ^2 (the variance in ability).

Differences in X In order to test proposition 2, that the variance in pay changes is decreasing in the degree of asymmetry in the market but more so for movers, we must identify markets where information spreads more and less easily. My identification comes from differences in the degree to which communicating with outsiders is an important task for the job. The idea is that when a worker is interacting with clients or others outside the firm, information about that worker's quality can spread. For example, in professional service occupations (e.g., consultants, and lawyers in law firms), interacting with clients is common and can be quite valuable for career prospects in that many workers in these jobs are in-housed by their clients.²⁶ In addition, the marketing literature emphasizes that in the professional services sector, long-term relationships between service providers and clients are common, in large part because clients can generate referrals and increase outside credentials (see for example Halinen 1997).

To measure the importance of this communication task I use O*NET 4.0 (formerly the Dictionary of Occupation Titles), which contains information on the task content of jobs.²⁷ The task of interest for my purposes is "Com-

²⁵Mean differences are obtained by regressing a characteristics on indicators for entry unemployment rate quartile. Variance differences are estimated by regressing the squared residuals from the previous regression on the same set of indicators. Standard errors are clustered by entry year.

²⁶This can be seen, for example, in top-consultant.com (2007), a report on retention in the consulting industry, in which they conducted a survey of over 700 consultants in 140 firms and asked questions about turnover behavior.

²⁷Available from the National Center for O*NET Development. I use O*NET 4.0, the

communicating with Persons Outside Organization." All occupations are rated in both the level and relative importance of this work activity on a scale from 0 to 7. To gain some intuition on this measure, appendix table B.2 shows the average rating for somewhat aggregated 3-digit 1970 Census Occupation codes, sorted from highest to lowest in the relative importance of communicating outside the firm, with the level of communication also shown.²⁸ As can be seen, salespeople and service workers (such as counselors, lawyers, teachers, and health care workers) are towards the top, while scientists, engineers, operatives, craftsmen and cleaners are towards the bottom. In this paper, I divide occupations into four groups representing quartiles of this measure, hereafter called communication groups, and use these as my source of variation in X (the degree of asymmetry). I use the measure of relative importance, rather than level, because I believe that is closest in spirit to X . However, these measures are highly correlated and results are similar when the level of importance is used.²⁹

Before moving on to my estimation strategy, I provide some suggestive evidence that occupations with more communication outside the firm look less like asymmetric information markets. The asymmetric information model pre-

earliest version available in a usable database. These data, from 2002, more closely match my sample period (1979-2000) than more recent updates.

²⁸O*NET categorizes occupations using 2000 SOC codes. To merge O*NET values in with the NLSY data, I must create a crosswalk from 2000 SOC codes to 1970 Census Occupation codes. I use the 2000 decennial census extract from IPUMS to match 2000 SOC codes to Census 1990 codes. I then use the 1970 decennial census extract from IPUMS to match Census 1990 codes to Census 1970 codes. Before these steps, I must aggregate some SOC codes in the O*NET data, assigning the average O*NET values of the disaggregated groups, to match the level of aggregation found in the 2000 decennial. I must also impute some values that are missing in O*NET, such as the "All Other" categories, to group-level averages. For example, SOC Code 119199, "Managers, all other" is missing in O*NET, so I assign to that group the average O*NET values of all the subcategories of manager. Finally, I match by hand some Census 1990 codes that do not show up in the 2000 decennial (but do in the 1970 decennial) to a 2000 SOC code. For 1990 census codes that match to more than one 2000 SOC code, I keep the modal one. More details are available upon request. I am able to match about 91% of the jobs in the NLSY.

²⁹In a previous version of this paper, I compared professional service workers to professional non-service workers, arguing that service workers have more outside communication and therefore should be in more symmetric markets. Results using this strategy are very similar and are available upon request.

dicts that lower ability workers are more likely to leave whereas the symmetric learning model predicts turnover is random. Therefore, I look at whether workers who leave jobs with higher levels of outside communication appear less negatively selected (when compared with stayers) than workers who leave jobs with less outside communication.

Table 2 summarizes regressions yielding differences in characteristics between movers and stayers, within communication group, relative to the lowest quartile group.³⁰ The table also shows the main effect for moving and shows movers are indeed negatively selected; they have lower AFQT scores, years of school, and tenure in the year before they moved, relative to stayers. This is consistent with the asymmetric information model and also with Schönberg (2007) who finds a roughly similar effect for AFQT in these same data. However, I also find movers have larger log wage changes between periods, consistent with work by Topel and Ward (1992), and many others, who show young workers tend to move to better opportunities and obtain higher pay increases upon moving.

More importantly, the coefficients on moving interacted with communication group are for the most part positive and significant. That is, the mover effects for workers in jobs with higher outside communication are less negative than workers who left jobs in lower communication groups. In addition, the final column of table 2 looks at the probability of being laid off or fired conditional on moving between periods. It reveals that workers in higher communication groups are less likely to leave their jobs involuntarily, another indication that they are less negatively selected. This finding that movers from markets with less outside communication are more negatively selected reinforces my hypothesis that they are in more markets where information is more asymmetric.

³⁰A characteristic is regressed on indicators for the communication group, a dummy for whether the worker moved (regardless of whether the move was to a different communication group), and interactions. Standard errors are clustered by 3-digit 1970 Census occupation, since that is the level of variation driving the communication groups.

4.1 Methodology

First, I estimate a pay-change regression of the form specified in equation 12 and obtain the residuals.

$$\begin{aligned}
 Lwage_{ijt} - Lwage_{ij,t-1} = & \alpha_0 + \alpha_1 I_{ijt}^{\text{stay}} + \alpha_2 I_{ij}^{\text{unemp}} + \alpha_3 I_{ij}^{\text{comm}} \\
 & + \alpha_4 [I_{ijt}^{\text{stay}} * I_{ij}^{\text{unemp}}] + \alpha_5 [I_{ijt}^{\text{stay}} * I_{ij}^{\text{comm}}] \\
 & + \alpha_6 Y_i + \alpha_7 Z_{ijt} + \alpha_8 [I_{it}^{\text{comm}} * I_{it-1}^{\text{comm}}] + \varepsilon_{ijt}
 \end{aligned} \tag{12}$$

For a worker, i , holding job, j , in year $t - 1$ (but not necessarily in year t), I regress log wage change on an indicator for whether the worker stayed in job j between periods t and $t - 1$ (I_{ijt}^{stay}), vectors of entry unemployment rate group fixed effects (I_{ij}^{unemp}) and communication group fixed effects (I_{ij}^{comm}) where both refer to the job held in $t - 1$, interactions of each of these with whether the worker stayed, a main effect for staying and a set of control variables.³¹ Y_i is a vector of time-invariant controls, while Z_{ijt} is a vector of time-varying controls.³² Lastly, I control for interactions of communication group in year t with communication group in year $t - 1$. Appendix table B3 shows summary statistics.

The goal with the first-stage equation is to condition on characteristics the employer can observe at the start of the employment relationship. The residuals obtained from this regression arguably yield a better empirical analogue to the spirit of the model (than simply using log wage changes), where ability,

³¹In my sample a worker may not hold more than one job at a time, but may show up with a number of jobs over the entire sample period. Note values for I^{unemp} and I^{comm} refer to the current job and remain the same regardless of whether the worker moves between $t - 1$ and t , and regardless of where the worker moved (e.g., the worker could move to a job in a different communication group, or not).

³² X_i includes black, Hispanic and education dummy variables. Z_{ijt} contains a quadratic in age in year $t - 1$, dummy variables for years tenure (rounded) in both t and $t - 1$, the square of the change in tenure, final tenure in the $t - 1$ job if the worker left otherwise 0, indicators for geographic region and urban status, marital status, actual labor market experience before starting the $t - 1$ job, calendar year indicators, occupation and industry dummies (both measured in 1970 one-digit level), indicators for having changed occupation or industry between period, an indicator equaling 1 if the worker was always observed post-transition in the dataset, and cumulative years of training.

μ , is unobservable.³³

I then measure the variance of these log wage-change residuals using the squared residuals from the first-stage regression. The variance of wage-change residuals for a particular group of workers, g , equals $E[\varepsilon_{ijt}^2|g] - (E[\varepsilon_{ijt}|g])^2$. The residuals are purged of group fixed-effects in the first-stage regression (where group is job-entry cohort or communication group, each interacted with staying), so $E[\varepsilon_{ijt}|g] = 0$. Thus the variance can be characterized by the ε_{ijt}^2 's.³⁴

To be precise, equation 13 shows the second-stage regression. This regression includes some of the controls from the first stage and is robust to the inclusion of others.³⁵ Because employer learning should be most prevalent early in a worker's career, I also present results where this and the first-stage equation are restricted to workers with very low tenure (those who have been with their employer less than 2 years when the $t - 1$ wage is measured).

$$\begin{aligned} \varepsilon_{ijt}^2 = & \beta_0 + \beta_1 I_{ijt}^{\text{stay}} + \beta_2 I_{ij}^{\text{unemp}} + \beta_3 I_{ij}^{\text{comm}} + \beta_4 [I_{ijt}^{\text{stay}} * I_{ij}^{\text{unemp}}] + \beta_5 [I_{ijt}^{\text{stay}} * I_{ij}^{\text{comm}}] \\ & + \beta_6 Y_i + \beta_7 [I_{it}^{\text{comm}} * I_{it-1}^{\text{comm}}] + \nu_{ijt} \end{aligned} \quad (13)$$

The difference-in-differences estimates, β_4' , correspond to dd_σ from equation 10, while the β_5' correspond to dd_X from equation 11. Recall from the previous

³³One might worry about variables the employer can observe but I, the econometrician, cannot, especially if they differ systematically across job-entry cohorts or communication groups. This is not a problem as long as these variables are observed by all employers, in which case they are absorbed in the group fixed effects (for job entry cohorts and communication groups, each interacted with staying) in the first stage. For example, that workers who enter firms in recessions are, on average, lower ability, is absorbed in the job entry unemployment rate quartile fixed effects.

³⁴Using squared residuals allows me to exploit all variation in the micro data by keeping the unit of observation an individual-employer-year. However, results are similar when I collapse the data to the job-entry cohort*stay or communication group*stay level and estimate standard deviations. Also, because the expectation of squared residuals is particularly sensitive to outliers, I have also looked at the inter-quartile ranges for the residuals and obtained very similar results.

³⁵It seems particularly important to allow for the fact that different education levels and races will have different pay dispersions, hence the inclusion of Y_i . In addition, the interactions of communication group in year $t - 1$ with that in year t , allows me to control for the average spread of wage changes for each occupation combination.

section, a finding of $dd_\sigma < 0$ and $dd_X < 0$ are consistent with asymmetric information. The former implies that the decreased ability variance seen in cohorts matching to a firm in a recession compared to those matching in a boom has a larger compressing effect on the distribution of pay changes for stayers relative to movers. The latter implies that the decrease in the degree of asymmetry seen in jobs with more outside communication, relative to less, has a larger widening effect on the wage change distribution of movers compared to that of stayers.

As noted above, the spread of wage changes may differ for movers, relative to stayers for reasons outside the scope of the model. Furthermore, there is surely non-random selection into job-entry cohorts and communication groups, so we might expect the spread of wage changes to differ as well. My difference-in-differences strategy allows me to control for all of these factors with fixed effects for each group (represented by γ^{stay} , γ^σ , and, γ^X in equation 5). Importantly, the key identifying assumption is that there is no differential selection in turnover across groups outside that specified in the model, i.e., γ^{stay} is the same across job-entry cohorts and is the same across communication groups. Below I provide evidence that similar turnover is a reasonable assumption in this case.

One way in which job-entry or communication groups could differ is in their human capital investments, an issue left unaddressed in the model. The impact of group-level differences in human capital accumulation on the wage will be absorbed in the group fixed effects, however differential accumulation within groups across movers and stayers would not. I therefore control for self-reported investments in training in the first stage (results are very similar when I also include training in the second stage regression).³⁶

One final note on clustering. For the job cohorts analysis, the right level of clustering is the entry year. For the communication group analysis, clustering at the 3-digit 1970 Census occupation level makes most sense. One could also

³⁶Each year, respondents were asked if they participated in a training program, though the definition of training program changes across years. Follow the same approach as Altonji and Pierret (2001), I create a cumulative effect by summing across years then control for this effect in the first-stage regressions.

make the case for clustering at the individual level since workers can appear in multiple jobs over the sample period. As a general rule of thumb, clustering at the most aggregated level is most conservative. I thus cluster at the entry year level. This is indeed the more conservative choice since I find that statistical significance always weakly increases when clustering on three-digit occupation or at the individual level.

5 Results

In this section, I test propositions 1 and 2, and address the main identifying assumption that turnover rates do not differ systematically across groups, except via mechanisms specified in the model. I also use these estimates to provide two approximations for X , the degree of asymmetry, in the economy as a whole. Proposition 1 says that under asymmetric information, the variance in wage changes of stayers responds *more* to different variances in ability than does the variance of wage changes for movers. This is because the wage offers of outside firms are less linked to individual workers' abilities. I test this by exploiting variation in ability distributions across job-entry cohorts, where those who entered in worse economies have a lower variance in ability – likely due to their negative selection since a smaller share of them were voluntary quitters on their previous job. Proposition 2 says that under asymmetric information, the variance in wage changes of stayers responds *less* to different levels of asymmetry than does the variance of wage changes for movers. This is because the degree of asymmetry impacts outside firm expectations, a larger component of outside wage offers than incumbent wage offers. The empirical implementation is to compare jobs with different degrees of communication outside the firm (using measures from O*NET), hypothesizing that more information about worker quality is transmitted to outside firms in jobs with higher communication levels.

5.1 Propositions 1 and 2

Table 3 summarizes the core set of results for this paper, regressions of the form specified in equation 13. The dependent variable is squared residuals from the log wage change regression and the goal is to estimate differences in the variance in wage changes. Recall the interactions of staying at the firm and job-entry cohort or communication group (β'_4 and β'_5) test propositions 1 and 2 in the model. I report these interactions, as well as the main effect for staying, and highlight in particular the highest unemployment rate and communication groups, relative to the lowest, since these should yield the sharpest contrast.

Note first that the coefficient on staying is always negative and significant. This is expected since there are many match-specific factors outside the scope of the model (γ^{stay}) which remain constant when a worker stays at a firm (for example, compensating differentials) and vary when workers move firms. These are important to control for through the inclusion of this main effect for staying.

The first two columns show results for the full sample and the low tenure sample, respectively. Here we see that the difference-in-differences estimate for the worst economy entry cohorts, relative to the best, are negative and statistically significant at the 1% level. The magnitude is larger in the low tenure sample, where employer learning should be more important. I find, in the low tenure sample, that the variance in log pay changes is 0.056 more compressed for stayers, relative to movers, among those entering in the worst, compared to the best, economy. The coefficients on the other entry cohort interactions fall in magnitude, basically as would be expected, given these groups should be more similar to the lowest unemployment rate group. Wage changes are thus lower variance for workers who enter firms in a bad economy than a good one and stay, relative to this comparison among movers. Since these entrants are also lower variance in ability, this finding is consistent with proposition 1, that differences in worker ability distributions are more strongly reflected in the wage change distributions of stayers, relative to movers.

The first two columns of table 3 also present results consistent with proposition 2. The difference-in-difference estimates for the highest communication

group, relative to the lowest, are negative and, in the low tenure case, statistically significant at the 5% level.³⁷ I find, in the low tenure sample, that the variance in log pay changes is 0.083 more compressed for stayers, relative to movers, among those in the highest, compared to the lowest, communication group. Moving to lower communication groups also reduces the magnitude of the coefficients in an expected way, given these groups are more similar to the lowest group. Wage changes are thus more compressed for stayers, relative to movers, in the more symmetric market. Or, moving to a more symmetric market, widens the wage change distribution for movers more than for stayers, supporting proposition 2.

Table 3 also provides estimates separately by education, a BA or more in the second set of columns and less than a BA in the third set of columns. This is of interest since Schönberg finds some evidence in favor of asymmetric learning among college graduates but not among high school graduates. My findings here are inconclusive.

The estimates of dd_X (communication groups) are larger in magnitude, negative and statistically significant (consistent with proposition 2) for the high education group, but for the low education group, though magnitudes are sizeable, estimates are not statistically significant. These estimates thus lend some support to the Schönberg finding, that asymmetric information may be present in the college graduates market. On the other hand, the estimates of dd_σ are large in magnitude, negative and statistically significant (consistent with proposition 1) in the low education group, but not in the high one (though the magnitude in the low tenure case is sizeable). However, though not shown, I find that variances in the characteristics shown in table 1 do not differ significantly across groups among the high educated. Therefore we would not expect the same predictions for the dispersion of wage changes. This particular identification strategy is therefore not well-suited to identifying differences in the degree of asymmetry across education groups.

³⁷When regressions are clustered by 3-digit occupation (the level of variation underlying the communication groups) this coefficient in the low tenure sample is statistically significant at the 1% level.

I would therefore note my findings are broadly consistent with the previous literature, since most papers find support for asymmetric information. Schönberg, on the other hand, fails to reject the symmetric learning model overall, though has some weak evidence in favor among college graduates. Several factors could be contributing to the differences in our findings. First, as I noted in the introduction, the AFQT approach requires several strong assumptions that my approach does not. This could easily yield different conclusions. A particularly important assumption in Schönberg's analysis is that turnover is only driven by the mechanisms specified in her model. She does not account for the impacts of compensating differentials, search theory, match quality, and ability sorting (among others) on the turnover decision. My model also abstracts away from these factors, but I have shown that I can empirically account for them via γ^{stay} , the fixed effect for staying. I rely on the weaker assumption that these factors do not differentially impact turnover across job entry groups or communication groups. An additional assumption of the AFQT approach, which Schönberg notes, is that it is best suited to test whether firms learn about cognitive ability of workers, while my approach does not have such a restriction. It could therefore be the case that asymmetric learning about non-cognitive skills is most important. Lastly, I would say that though Schönberg cannot reject the symmetric learning model, her standard errors are large enough that she cannot rule out a large role for asymmetric learning either.

More generally, the AFQT approach can only evaluate employer learning about the component of worker productivity measured before workers begin their labor market experience (since the AFQT was administered in the NLSY when most respondents were teenagers). Lange (2007) finds, using a symmetric learning model, that employers learn quite quickly about productivity differences that exist at the beginning of the career. Therefore, we might not expect much scope for asymmetric learning about these initial productivity differences. However, Kahn and Lange (2012) find, using proprietary data on pay and repeat performance evaluations, that productivity continues to evolve in unpredictable ways as workers age and show that employer learning

is important at all stages of the life-cycle. It is therefore plausible that asymmetric learning is also an important mechanism in the labor market, even if asymmetric learning about productivity differences contained in AFQT score is not.

5.2 Turnover Rates

As highlighted above, my empirical strategy identifies dd_σ and dd_X based on the interactions of job-entry group or communication group and staying at the firm. Thus, I cannot have differential mobility across groups, outside that which is specifically modeled. Looking directly at turnover rates across these groups is tempting, since that will give some indication. However, it is easy to show that in my asymmetric information model, turnover is higher among workers with lower variance ability and in markets where learning less asymmetric. Thus, we should expect to see higher turnover rates among those who matched to firms in the worst economies and those in the highest communication group.

Table 4 summarizes regression results where I estimate the probability of staying at the firm on the job-entry and communication group main effects. Columns labeled "1" include the same set of controls used in the wage change equations (except the second period variables). Though not statistically significant, these coefficients on the highest unemployment rate group are negative in both samples, consistent with the model.³⁸ I will use the estimates in my second approximation of X , below. The coefficients on the highest communication group are positive, though very close to zero. Note, I do not view this as a direct test of my model since I have not shown that the derivatives of the probability of staying with respect to σ^2 and X are empirically identified – recall the importance of γ^{stay} , γ^σ , and γ^X in the implementation section.

Columns labeled "2" in table 4 additionally control for age-adjusted AFQT and log wage in $t - 1$. The hope is that these variables help to pick up the mobility effects specified in the model. AFQT can give some indication of

³⁸The national unemployment rate effect is consistent with Bowlus (1995), who also finds using these data that jobs are shorter in duration when the match begins in a recession.

unobserved ability and should thus control for the negative selection of movers predicted in the model. Wages in the model incorporate incumbent and outside firm expectations, which are both a function of the degree of asymmetry and worker ability. The relationship is obviously rough at best. Still, we see in most instances, the magnitudes on the coefficients of interest reduce with the inclusion of these controls.

Given the lack of statistical significance and the relatively small magnitudes (compared to the sample mean for staying of 65%) presented in table 4, I find little evidence that is inconsistent with my identifying assumption.³⁹ Turnover in this sample is obviously important to account for, beyond the mechanisms of adverse selection and the random disutility shock specified in my model, especially because workers in this sample are young and move jobs often (see for example Topel and Ward 1992). In fact, the average worker in my sample has held slightly more than 6 jobs. However, because workers in my sample are young and move jobs often, my identifying assumption that workers across job entry and communication groups do not face differential turnover decisions (beyond those specified in the model) may be less worrisome.

5.3 The Size of X

Table 3 provides evidence consistent with propositions 1 and 2, and therefore consistent with asymmetric learning being prevalent in the labor market. The question, then, is how prevalent, or how important? To answer that question, we would like to estimate X , the degree of asymmetry. In this subsection, I provide two rough approximations of X , implied by my model. Both have advantages and disadvantages, discussed in more detail below, and are meant to serve as a guide for interpreting the coefficients in table 3. Interestingly, though, methods 1 and 2 arrive at roughly similar values for X . Both are described in detail in appendix A.4.

³⁹I can also control for the reason the worker left the previous job (for those who moved between $t - 1$ and t , and otherwise 0) in both the first and second stage regressions. This could be one way to control for differential match quality across groups. I find very similar results with the inclusion of these controls.

Method #1 One way to solve for X is by taking the ratio of ∂_σ and ∂_X , the derivatives of $V^{stay} - V^{move}$ with respect to σ^2 and X , respectively (equations 11 and 12 in the appendix) and solving. It is easy to show that $\frac{\partial_\sigma}{\partial_X} = \frac{X}{\sigma^2}$, which yields a nice, tractable equation for finding X . In section 3, above, I showed how the empirical analogues, dd_σ and dd_X , relate to these derivatives. Substituting in and solving for X , we get equation 14.

$$X = \frac{dd_\sigma \sigma^2 \Delta X}{dd_X \Delta \sigma^2} \quad (14)$$

We have empirical estimates of dd_σ and dd_X . Specifically, I take β_4 from the highest unemployment rate group and β_5 from the highest communication group (see regression equation 13) estimated on the low tenure sample. However we do not have estimates of the other terms. In particular, we would like to know the relative sizes of variation in σ^2 and X (normed by σ^2) induced by moving from highest to lowest national unemployment rate entry cohorts and communication groups, respectively. Unfortunately, my measure of the degree of asymmetry is categorical – the importance of communication outside the firm ranges from 0 to 7. Thus I cannot measure which exercise results in a larger shift. Instead, I provide a plausible range for X by varying $\frac{\sigma^2 \Delta X}{\Delta \sigma^2}$ from 0.8 to 1.2. That is, I vary the relative magnitudes from a shift across entry cohorts having 80% the impact on σ^2 that a shift across communication groups has on X (normed by σ^2) to the reverse.

Table 5 shows these approximations of X . I find that the degree of asymmetry ranges from 0.54 to 0.81. This range of estimates implies a large role for asymmetric learning. However, the ad hoc calibration of $\frac{\sigma^2 \Delta X}{\Delta \sigma^2}$ is somewhat unsatisfying as it produces a fairly wide range of estimates. I therefore provide an additional approximation of X , below, which uses a series of equations from my model to solve for X and the additional unknowns.

Method #2 This method uses the difference-in-differences equation for dd_σ (derived by combining equations 6-9, above). This is a function of four unknowns: $\sigma_H^2, \sigma_L^2, X$ and $\bar{\theta}$. I therefore need 3 other equations or identifying

assumptions to obtain a solution for X . I start with the probability that a group of workers stays with the incumbent firm, which is empirically estimable and is analytically a function of $X, \bar{\theta}$, and the relevant σ^2 . Using two groups of workers, the high unemployment rate group and the low unemployment rate group, I can easily estimate the probability that each group stays with their incumbent, and have done so in table 4. This yields two more equations to be used in identifying X . I then employ one more empirical assumption, explained next, to pin down $\Delta\sigma^2 = \sigma_L^2 - \sigma_H^2$, leaving me with 4 equations and 4 unknowns. Equation 18 in the appendix uses these to solve for X .⁴⁰

To pin down $\Delta\sigma^2$, the difference in the variance in ability between the high- and low-variance groups, I take advantage of AFQT score. As discussed above, AFQT is an aptitude test score that has been shown to be correlated with productivity.⁴¹ I showed in table 1 that the highest unemployment rate entry cohort has a substantially smaller variance in AFQT score, compared to the lowest unemployment rate entry cohort and I use that difference here.⁴²

The last column of table 5 shows these estimates of X , again using the low tenure sample.⁴³ I find X equals 0.69, exactly in the middle of the previous

⁴⁰Note, I could employ an analogous strategy based on dd_X , yielding yet a third estimate of X . However, I do not believe I have plausible empirical identification of ΔX (that is the change in X induced by moving from one communication group to another), leaving me without a crucial fourth equation to pin down the parameters. So I do not explore this option.

⁴¹The score has been normed to be mean 0, standard deviation 1 over the whole sample, thus the variance of AFQT score is meaningless. Instead, I here use the *difference* in variances between national unemployment rate entry cohorts, which should be more informative.

⁴²AFQT represents only differences in ability that exist before entry into the labor market. If these initial differences fan out over time, then I have underestimated the magnitude of $\Delta\sigma^2$. From equation 18 in the appendix I can show that a smaller magnitude of $\Delta\sigma^2$ results in an overestimate of X . However, this is simply an approximation, meant more as a way of interpreting the magnitudes of my main coefficients in table 3.

⁴³Specifically, for dd_σ , I use the coefficient on the highest unemployment rate group, in the low-tenure sample in table 3. I estimate the probability of staying with the incumbent in the low tenure sample (0.62) and use that for the lowest national unemployment rate group. For the highest national unemployment rate group, I add to the sample mean the coefficient for this group from table 4 (resulting in 0.59 probability of staying), thus holding constant all other control variables. For $\Delta\sigma^2$, I follow the same methodology as in table 1 and find a difference of -0.06 in the low-tenure sample across highest and lowest national unemployment rate groups.

range of estimates.

Discussion Translating these estimates into differential speeds of employer learning provides a useful interpretation. Lange (2007) defines the speed of employer learning as the relative weight employers place on the first new signal of ability. Though his is a multi-period model and mine is a two-period model, we could roughly borrow his definition and in this case, the speed of employer learning equals $1 - X$. As Lange points out, the speed of employer learning is nicely interpreted as the reduction in expectation errors over worker productivity in one period of learning. In my case, reduction in expectation errors is relative to the incumbent whose initial expectation errors decline by 100%, by assumption.

The second row of table 5 shows that the speed of employer learning ranges from 0.19 to 0.46, in method #1, and equal to 0.31 in method #2. Thus expectation errors over initial productivity decline by roughly 31% of their decline in incumbents. In other words, I estimate that outside firms learn only roughly a third of what incumbents learn. Thus, using either method, I find a sizeable role for asymmetric information.

Roughly speaking, we could extend the learning process specified in my model to a multi-period model where outside firms receive a signal of worker quality each period that contains the same noise structure. As shown in the third row, my estimates would imply that in 5 periods outside firms have learned 54%-81% of what incumbents learned in the first period. In other words, after 5 periods, outside firms have learned roughly 70% (from method #2) of what incumbents learned in the first period.

The two methods produce remarkable consistent estimates despite their different identifying assumptions. Method #2 has the advantage of being able to pin down one estimate of X , as opposed to the range produced in method #1. It has the disadvantage that I need to use equations from the model that I have not shown to be empirically identified. It also requires a further assumption on $\Delta\sigma^2$, for which I use AFQT score. Method #1 is nice in that it only uses empirical estimates which I have shown are identified (dd_σ and

dd_X) and it takes advantage of empirical variation in both σ^2 and X , while method #2 only exploits variation in σ^2 . However, it requires an assumption about $\frac{\sigma^2 \Delta X}{\Delta \sigma^2}$, for which I supply a plausible range of values. I offer both sets of estimates as a way of interpreting the main coefficients of interest in table 3, though I caution against taking them too literally.

6 Conclusion

This paper seeks to establish whether asymmetric employer learning is prevalent in the labor market. I derive a model which embeds both symmetric and asymmetric learning and develop empirical predictions about the variance in pay changes to test for the prevalence of asymmetric learning. I then test the model using two identification strategies. The first exploits differences in job-entry cohorts to test comparative statics over the variance in ability. The second tests comparative statics over the degree of asymmetry in the market, using variation across occupations in the degree to which communicating outside one's firm is an important task. Both predictions are supported in the data. Finally, I show that the magnitude of the degree of asymmetry is economically important. I find that outside firms have sizeable expectation errors over worker ability, relative to the incumbent. In one period, outside firms reduce expectation errors by roughly a third of the incumbent's reduction.

These results are consistent with most of the previous empirical work in this area, though there have only been a few papers. Despite examining very different samples, Gibbons and Katz (1991), DeVaro and Waldman (2012), and Acemoglu and Pischke (1998) all find strong support for asymmetric information.⁴⁴ Pinkston (2009) and Zhang (2007) use the NLSY (comparing the return to AFQT score across tenure versus experience) and find evidence consistent with asymmetric learning. However, as previously discussed, Schönberg (2007) does not. Besides differences in assumptions, my methodology allows

⁴⁴Gibbons and Katz use the CPS Displaced Workers Supplement, DeVaro and Waldman use proprietary data from a single firm, and Acemoglu and Pischke provide evidence from German apprenticeship firms.

me to test for asymmetric learning about both cognitive and non-cognitive ability and about ability as it evolves over time, while the use of AFQT score is best thought of as a correlate of cognitive ability measured before labor market entry. A fruitful avenue for future research would be to investigate the relative importance of learning about each of these components of ability.

The previous literature and the current paper, therefore, largely support the existence of asymmetric information and its prevalence. These findings then beg the question: what is the impact of asymmetric information on worker mobility, investments in human capital and allocating workers to jobs and tasks? Furthermore, what policies have firms already implemented to take advantage of their information rents? The model presented in this paper is useful in that it provides a methodology to study learning in a variety of settings without requiring particularly demanding datasets. Extensions could clarify thinking about the effects asymmetric information has on contracting, promotion policy, human capital investment and turnover behavior and perhaps policies that could alleviate inefficiencies.

References

- [1] Acemoglu, D. and J.-S. Pischke (1998): "Why Do Firms Train: Theory and Evidence," *Quarterly Journal of Economics*, 113:79-119.
- [2] Altonji, J.G. and C.R. Pierret (2001): "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 113:79-119.
- [3] Bernhardt, D. (1995), "Strategic Promotion and Compensation," *Review of Economic Studies*, 62, pp. 315-339.
- [4] Bowlus, A. (1995): "Matching Workers and Jobs: Cyclical Fluctuations in Match Quality," *Journal of Labor Economics*, 13:335-350.
- [5] Chang, C. and Y. Wang (1995): "A Framework for Understanding Differences in Labor Turnover and Human Capital Investments," *Journal of Economic Behavior and Organization*, XXVIII: 91-105.

- [6] Chang, C. and Y. Wang (1996), "Human Capital Investment Under Asymmetric Information: The Pigovian Conjecture Revisited," *Journal of Labor Economics*, 14, pp. 555-570.
- [7] DeGroot, M. H. (2004): *Optimal Statistical Decisions*, John Wiley & Sons, Inc., Hoboken, NJ.
- [8] DeVaro, J. and M. Waldman (2012): "The Signaling Role of Promotions: Further Theory and Empirical Evidence," *Journal of Labor Economics*, 30(1), pp. 91-147.
- [9] Farber, H.S. (1999): "Mobility and Stability: The Dynamics of Job Change in Labor Markets," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, vol. 3, North-Holland, pp. 2439-2483.
- [10] Farber, H.S. and R. Gibbons (1996): "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 111:1007-1047.
- [11] Gibbons, R. and L.F. Katz (1991): "Layoffs and Lemons," *Journal of Labor Economics*, 9:351-380.
- [12] Gibbons, R. and M. Waldman (1999): "A Theory of Wage and Promotion Dynamics inside Firms," *Quarterly Journal of Economics*, 114:1321-1358.
- [13] Gibbons, R., L.F. Katz, T. Lemieux, D. Parent (2005): "Comparative Advantage, Learning and Sectoral Wage Determination," *Journal of Labor Economics*, Vol. 23(4): pp. 681-724.
- [14] Greenwald, B.C. (1986): "Adverse Selection in the Labor Market," *Review of Economic Studies*, 53:325-347.
- [15] Halinen, A. (1997): "Relationship Marketing in Professional Services." Routledge: London.
- [16] Kahn, L.B. (2010): "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy," *Labour Economics*, 17(2).

- [17] Kahn, L. B. and F. Lange (2012): "Employer Learning, Productivity and the Earnings Distribution: Evidence from Performance Measures," Yale University, mimeo.
- [18] Katz, E. and A. Ziderman (1990), "Investment in General Training: The Role of Information and Labour Mobility," *Economic Journal*, 100, pp. 1147-1158.
- [19] Lange, F. (2007): "The Speed of Employer Learning." *Journal of Labor Economics*, 25: 1-35.
- [20] Milgrom, P. and S. Oster (1987): "Job Discrimination, Market Forces, and the Invisibility Hypothesis," *Quarterly Journal of Economics*, 102, pp. 453-476.
- [21] Pinkston, J.C. (2009): "A Model of Asymmetric Employer Learning with Testable Implications," *Review of Economic Studies*, January.
- [22] Prendergast, C. (1992): "Career Development and Specific Human Capital Collection," *Journal of the Japanese and International Economics*, VI: 207-227.
- [23] Schönberg, U. (2007): "Testing for Asymmetric Employer Learning" *Journal of Labor Economics*, 25, pp. 651-692.
- [24] top-consultant.com (2007), "Retention Report 2007: A Key Challenge for the Management Consultancy Sector."
- [25] Topel, R. and M. Ward (1992): "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 107: 441-479.
- [26] von Wachter, T. and S. Bender (2006): "In the Right Place at the Wrong Time - The Role of Firms and Luck in Young Workers' Careers," *American Economic Review* 96: 1679-1705.
- [27] Waldman, M. (1984): "Job Assignment, Signalling and Efficiency," *RAND Journal of Economics*, 15:255-267.

- [28] Waldman, M. (1990): "Up-or-Out Contracts: A Signaling Perspective," *Journal of Labor Economics*, 8, pp. 230-250.
- [29] Zabochnik, J. and D. Bernhardt (2001): "Corporate Tournaments, Human Capital Acquisition, and the Firm Size-Wage Relation," *Review of Economic Studies*, 68, pp. 693-716.
- [30] Zhang, Y. (2007): "Employer Learning under Asymmetric Information: The Role of Job Mobility," Indiana University-Purdue University Indianapolis, mimeo.

A Theory Appendix

A.1 Deriving $w^*(\mu, s)$ and $v^*(s)$

The worker chooses to stay at the firm if $w(\mu, s) - \theta > v(s)$. The probability of staying is expressed as follows, where the notation $F_\theta(\cdot)$ is introduced to represent the cumulative distribution function of θ (in this case the uniform distribution over the interval $[-\bar{\theta}, \bar{\theta}]$).

$$\Pr(\text{stay}|\mu, s) \equiv F_\theta(w(\mu, s) - v(s)) = \left\{ \begin{array}{ll} 0 & w(\mu, s) - v(s) < -\bar{\theta} \\ \frac{w(\mu, s) - v(s) + \bar{\theta}}{2\bar{\theta}} & w(\mu, s) - v(s) \in [-\bar{\theta}, \bar{\theta}] \\ 1 & w(\mu, s) - v(s) > \bar{\theta} \end{array} \right\}$$

The incumbent firm maximizes the expected profit for each ability type, μ , and realization of s , (i.e., the probability that the worker stays times the profit for that worker), taking outside firms' behavior as given. Denote the equilibrium outside wage offer as $v^*(s)$. The incumbent firm's problem and solution are as follows.

$$w^*(\mu, s) = \left\{ \begin{array}{ll} \max_{w(\mu, s)} F_\theta(w(\mu, s) - v^*(s))(\mu - w(\mu, s)) & \\ \frac{v^*(s) + \mu - \bar{\theta}}{2} & \mu - v^*(s) \in [-\bar{\theta}, 3\bar{\theta}] \\ v^*(s) + \bar{\theta} & \mu - v^*(s) > 3\bar{\theta} \end{array} \right\}$$

It is easy to show that the objective function is strictly quasi-concave when $w(\mu, s) - v(s) \in [-\bar{\theta}, \bar{\theta}]$. Thus the interior solution is indeed a maximum and is unique. We also have a corner solution for large μ . Below, I discuss how I deal with μ outside the middle range.

Outside firms set a wage schedule, $v(s)$, equal to expected productivity conditional on leaving the firm and on s , taking incumbent firm behavior ($w^*(\mu, s)$), as given. Bayes' rule yields the following expression, where $\pi(\cdot)$ is

the posterior pdf of μ conditional on s .

$$\begin{aligned}
v^*(s) &= E(\mu|\text{move}) = \frac{\int_{-\infty}^{\infty} \mu * (1 - F_{\theta}(w^*(\mu, s) - v(s)) * \pi(\mu|s) d\mu}{1 - F_{\theta}(w^*(\mu, s) - v(s))} \\
&= \frac{\int_{-\infty}^{v^*(s)-\bar{\theta}} \mu * 1 * \pi(\mu|s) d\mu + \int_{v^*(s)-\bar{\theta}}^{v^*(s)+3\bar{\theta}} \mu * \frac{-w^*(\mu, s)+v(s)+\bar{\theta}}{2\bar{\theta}} \pi(\mu|s) d\mu}{\int_{-\infty}^{v^*(s)-\bar{\theta}} 1 * \pi(\mu|s) d\mu + \int_{v^*(s)-\bar{\theta}}^{v^*(s)+3\bar{\theta}} \frac{-w^*(\mu, s)+v(s)+\bar{\theta}}{2\bar{\theta}} \pi(\mu|s) d\mu} \tag{1}
\end{aligned}$$

In order to obtain a closed-form solution, let us only focus on the portion of the distribution s.t. $\mu|s \in [v^*(s) - \bar{\theta}, v^*(s) + 3\bar{\theta}]$. For now, let us assume that the entire distribution, $\pi(\cdot)$, falls in this interval. In this case, I can replace equation 1 with equation 2. In other words I assume that all types have some probability of leaving that is strictly between 0 and 1. When the variance over the distribution of $\mu|s$ is small, relative to the variance of θ , this will be true for almost all workers. I formalize this condition in the following subsection. By plugging in for $\Pr(\text{stay}|\mu, s) = \frac{w(\mu, s) - v(s) + \bar{\theta}}{2\bar{\theta}}$ and $w^*(\mu, s) = \frac{v^*(s) + \mu - \bar{\theta}}{2}$, I solve for $v^*(s)$ as follows.

$$\begin{aligned}
v^*(s) &= \frac{\int_{-\infty}^{\infty} \mu * \frac{-w^*(\mu, s)+v(s)+\bar{\theta}}{2\bar{\theta}} \pi(\mu|s) d\mu}{\int_{-\infty}^{\infty} \frac{-w^*(\mu, s)+v(s)+\bar{\theta}}{2\bar{\theta}} \pi(\mu|s) d\mu} \tag{2} \\
&= \frac{E(\mu|s)(3\bar{\theta} + v^*(s)) - E(\mu^2|s)}{3\bar{\theta} + v^*(s) - E(\mu|s)} \\
0 &= (v^*(s))^2 + v^*(s)(3\bar{\theta} - 2E(\mu|s)) - E(\mu|s)3\bar{\theta} + E(\mu^2|s)
\end{aligned}$$

Using the quadratic formula, I solve for the roots of this equation. This is shown below where we know from DeGroot (2004) that the posterior distribution of μ , conditional on s , is normal with mean $mX + s(1 - X)$ and variance $X\sigma^2$.

$$v^*(s) = mX + s(1 - X) - \frac{3}{2}\bar{\theta} \pm \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}$$

The model yields two pure-strategy equilibria, both of which are intuitive.¹ The right root implies high period 2 wages and high turnover while the left root yields low period 2 wages and low turnover. If many workers leave the incumbent firm then outside firms attribute high ability to them and offer a high wage, thus the incumbent must offer high wages to keep workers. In contrast, if only a few workers leave the firm, outside firms attribute a low ability to them and offer them a low wage, allowing the incumbent to offer a low wage. However, the left root is not stable, since if another outside firm offers a slightly higher wage, that firm will attract workers AND earn a positive profit. The right root is stable since if another outside firm offers a slightly higher wage, that firm will attract all workers but make negative profit. Whereas, if an outside firm offers a slightly lower wage, that firm would attract no workers. Thus, I restrict attention to the right root for the rest of the analysis –though all results hold for the left root as well.

A.1.1 Placing a condition on $X\sigma^2$

To obtain a closed form solution to equation 1 above, I assumed that all values of the distribution of $\mu|s$ fall in the interval $[v^*(s) - \bar{\theta}, v^*(s) + 3\bar{\theta}]$. That is, I solved for $v^*(s)$ assuming that the probability of moving was strictly between 0 and 1 for all types of workers. This will be nearly correct if the distribution of θ is much higher variant than the distribution of $\mu|s$. If so then even a very high- μ worker could be induced to leave due to a negative utility shock and a low- μ worker could be induced to stay due to a positive utility shock.

I now formalize this notion by deriving conditions under which a large portion of the distribution of $\mu|s$ falls in the above interval. I define "large portion" as 99% of the distribution (or within ± 3 standard deviations since $\mu|s$ is normally distributed). When this is the case, any error created by assuming 100% is minimal and will not affect the comparative statics derived below.

Define $\Delta(\mu, s) = w^*(\mu, s) - v^*(s)$ and plug in the solutions from above to

¹Models of this form commonly yield multiple equilibria (see for example the following papers which focus on human capital acquisition under asymmetric information: Chang and Wang (1995), Prendergast (1992) and Acemoglu and Pischke (1998)).

get the following.

$$\Delta(\mu, s) = \frac{1}{2}(\mu - mX - s(1 - X) + \frac{1}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2})$$

For values of μ and s such that $-\bar{\theta} \leq \Delta(\mu, s) \leq \bar{\theta}$, the probability of moving is strictly greater than 0 and strictly less than 1 and we will therefore be in the desired interval. Let us define a new variable, $z = \mu - mX - s(1 - X)$. It is easy to show that $z \sim N(0, X\sigma^2)$. For ease of notation define $\sigma_z^2 = X\sigma^2$. Therefore we are interested in the probability that the following inequality holds. We now want to solve for the relationship between σ_z and $\bar{\theta}$ such that the interval $z \in [-3\sigma_z, 3\sigma_z]$ (3 standard deviations of z) always lies within the above inequality, or such that equation 3 holds.

$$\begin{aligned} -\frac{5}{2}\bar{\theta} + \sqrt{\frac{9}{4}\bar{\theta}^2 - \sigma_z^2} &\leq z \leq \frac{3}{2}\bar{\theta} + \sqrt{\frac{9}{4}\bar{\theta}^2 - \sigma_z^2} \\ -\frac{5}{2}\bar{\theta} + \sqrt{\frac{9}{4}\bar{\theta}^2 - \sigma_z^2} &\leq -3\sigma_z \text{ AND } 3\sigma_z \leq \frac{3}{2}\bar{\theta} + \sqrt{\frac{9}{4}\bar{\theta}^2 - \sigma_z^2} \end{aligned} \quad (3)$$

$$\text{iff } \sigma_z \leq \bar{\theta} \left(\frac{15 - \sqrt{65}}{20} \right) \approx \frac{7}{20}\bar{\theta} \quad (4)$$

It is easy to show that the solution is as given in equation 4. This is easily interpreted considering the standard deviation of θ (a uniformly distributed variable in the interval $[-\bar{\theta}, \bar{\theta}]$) is $\frac{\bar{\theta}}{\sqrt{3}}$. Therefore when the standard deviation over the posterior of μ conditional on s is approximately 60% of the standard deviation of θ or less, 99% of workers will have a probability of leaving strictly greater than 0 and less than 1. The remainder of the theory appendix assumes that 100% of works do.

Note also that if this condition holds, $\frac{9}{4}\bar{\theta}^2 > X\sigma^2$, and $v^*(s)$ has only real solutions.

A.2 Solving for the first-period wage, w_0

Assuming no discounting, the incumbent firm's ex-ante expected profits, Π , are as follows.

$$\Pi = E[\mu] - w_0 + E[(\mu - w(\mu, s))|stay]$$

Setting $\Pi = 0$ allows me to solve for w_0 as follows:

$$\begin{aligned} w_0 &= m + \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\mu - w(\mu, s)) \Pr(stay|\mu, s) \Pr(\mu|s) \Pr(s) d\mu ds}{\Pr(stay)} \\ &= m + \frac{1}{2} \left(\frac{X\sigma^2}{\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}} + \frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2} \right) \end{aligned}$$

Note that w_0 equals ex-ante expected ability, m , plus a positive constant. In period 2, the incumbent makes positive profits due to its information rents and must therefore pay the worker above average productivity in period 1 to balance this. It is easy to show that w_0 is increasing in X , the degree to which information in the market is asymmetric. When outside firms have less information, the period 2 quasi rent of the incumbent firm is larger.

A.3 Variance in wage changes

In this subsection, I derive the variance of pay changes for workers who stay at the firm (i.e., workers who accept offer w^*) and workers who move firms (accept v^*). I then prove propositions 1 and 2. Starting from the following expression for Δw^{stay} , I derive the variance.

$$\begin{aligned} \Delta w^{stay} &= \frac{1}{2}(\mu + mX + s(1 - X) - \frac{5}{2}\bar{\theta} + \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}) - w_0 \\ \text{Var}(\Delta w^{stay}) &= \frac{1}{4}(\text{Var}(\mu|stay) + (1 - X)^2\text{Var}(s|stay) + 2(1 - X)\text{Cov}(\mu, s|stay)) \end{aligned}$$

Taking each piece at a time, I first solve for $\text{Var}(\mu|stay)$. This uses the fact that the variance of the posterior distribution of μ , $\text{Var}(\mu|s)$, is a constant,

thanks to the normality assumptions. Using the law of iterated expectations, $Var(\mu|stay) = E[E(\mu^2|stay, s)|stay] - (E[E(\mu|stay, s)|stay])^2$. Therefore we have

$$\begin{aligned}
Var(\mu|stay) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu^2 \frac{\Pr(stay|\mu, s) \Pr(\mu|s) \Pr(s)}{\Pr(stay|s)} d\mu ds \\
&\quad - \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu \frac{\Pr(stay|\mu, s) \Pr(\mu|s) \Pr(s)}{\Pr(stay|s)} d\mu ds \right)^2 \\
&= \frac{E(\mu^3) - E[E(\mu^2|s)E(\mu|s)] - Var(\mu|s)E(\mu)}{\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}} - \frac{(Var(\mu|s))^2}{(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2})^2} + Var(\mu) \\
&= \frac{E((\mu - E(\mu|s))^3)}{\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}} + \sigma^2 - \frac{(X\sigma^2)^2}{(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2})^2}
\end{aligned}$$

But the $E((\mu - E(\mu|s))^3) = E[E((\mu - E(\mu|s))^3|s)]$ by the law of iterated expectations. The latter expression is the skew of the posterior distribution. Since the posterior is normally distributed, its skew is 0. Therefore equation 5 gives the variance of μ conditional on staying.

$$Var(\mu|stay) = \sigma^2 - \frac{(X\sigma^2)^2}{(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2})^2} \tag{5}$$

I now derive the variance of s conditional on staying.

$$\begin{aligned}
Var(s|stay) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s^2 \frac{\Pr(stay|\mu, s) \Pr(\mu|s) \Pr(s)}{\Pr(stay|s)} d\mu ds \\
&\quad - \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s \frac{\Pr(stay|\mu, s) \Pr(\mu|s) \Pr(s)}{\Pr(stay|s)} d\mu ds \right)^2 \\
&= \frac{\int_{-\infty}^{\infty} s^2 (\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}) \Pr(s) ds}{\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}} - \left(\frac{\int_{-\infty}^{\infty} s (\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}) \Pr(s) ds}{\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}} \right)^2 \\
&= E(s^2) - (E(s))^2
\end{aligned}$$

The conditional variance of s equals the unconditional variance of s . Because both outside employers and incumbent employers see s , the incumbent will

set wages so that, conditional on s , the worker is indifferent between staying and leaving. Since $s = \mu + \epsilon$ and $\mu \perp \epsilon$, the variance is just the sum of these variances, shown in equation 6.

$$Var(s|stay) = \frac{1}{1-X}\sigma^2 \quad (6)$$

Lastly, I solve for the covariance of s and μ conditional on staying.

$$\begin{aligned} Cov(\mu, s|stay) &= E(\mu * s|stay) - E(\mu|stay)E(s|stay) \\ &= \frac{Var(\mu|s)E(s)}{\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}} + E[E(\mu|s) * s] - \left(\frac{Var(\mu|s)E(s)}{\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}} + E(\mu)E(s) \right) \\ &= E[E(\mu|s) * s] - E(\mu)E(s) \\ &= \sigma^2 \end{aligned} \quad (7)$$

Plugging in equations 5-7 into the above formula for the variance in wage changes conditional on staying at the firm yields equation 8, the variance of wage changes of stayers.

$$Var(\Delta w^{stay}) = \sigma^2 + \frac{1}{4} \left(-3X\sigma^2 - \frac{(X\sigma^2)^2}{\left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}\right)^2} \right) \quad (8)$$

I next derive the variance in wage changes among movers. Starting from the following expression for Δw^{move} , I derive the variance.

$$\begin{aligned} \Delta w^{move} &= mX + s(1-X) - \frac{3}{2}\bar{\theta} + \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2} - w_0 \\ Var(\Delta w^{move}) &= (1-X)^2 Var(s|move) \end{aligned}$$

But above I showed that the variance of s conditional on staying equals the unconditional variance of s . Therefore the variance of s conditional on moving also equals the unconditional variance (since workers can only either stay or move). Since $var(s) = \frac{1}{1-X}\sigma^2$, we get equation 9:

$$Var(\Delta w^{move}) = (1-X)\sigma^2 \quad (9)$$

Propositions 1 and 2 involve the derivative of $Var(\Delta w^{stay}) - Var(\Delta w^{move})$ with respect to σ^2 and X , respectively. Therefore equation 10 subtracts equation 9 from equation 8 (where I use the abbreviated notation V^{stay} and V^{move} for the variances in wage changes of stayers and movers, respectively).

$$V^{stay} - V^{move} = \frac{1}{4} \left(X\sigma^2 - \frac{(X\sigma^2)^2}{\left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}\right)^2} \right) \quad (10)$$

Proposition 1 $\frac{\partial(V^{stay} - V^{move})}{\partial\sigma^2} > 0$

Proof. In order to prove proposition 1, we need to sign the derivative with respect to σ^2 which is shown in equation 11 and for ease of notation, written at ∂_σ .

$$\partial_\sigma = \frac{1}{4} X \left(1 - \frac{2X\sigma^2}{\left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}\right)^2} + \frac{(X\sigma^2)^2}{\sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2} \left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}\right)^3} \right) \quad (11)$$

It can be shown that as long as $\bar{\theta}$ is large compared to $X\sigma^2$ (the condition derived in section A.1.1. is more than sufficient), ∂_σ is always positive. ■

Proposition 2 $\frac{\partial(V^{stay} - V^{move})}{\partial r} > 0$

Proof. In order to prove proposition 2, we need to sign the derivative with respect to X , shown in equation 12 and for ease of notation, written as ∂_X . But if proposition 1 holds then ∂_X must also be greater than 0, since both X and σ^2 are positive.

$$\partial_X = \frac{1}{4} \sigma^2 \left(1 - \frac{2X\sigma^2}{\left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}\right)^2} + \frac{(X\sigma^2)^2}{\sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2} \left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma^2}\right)^3} \right) \quad (12)$$

■

A.4 Estimating X

The first method for empirically approximating X is to take the ratio $\frac{\partial_\sigma}{\partial_X}$, or the ratio of equations 11 and 12. It is easy to show that most of the terms

cancel out and we are left with equation 13.

$$\frac{\partial_\sigma}{\partial_X} = \frac{X}{\sigma^2} \quad (13)$$

In the empirical implementation I show that I can empirically estimate dd_σ and dd_X and that $dd_\sigma \approx \partial_\sigma \Delta\sigma^2$ and $dd_\sigma \approx \partial_X \Delta\sigma^2$. Substituting these expressions into equation 13 and solving for X yields equation 14.

$$X = \frac{dd_\sigma}{dd_X} \frac{\sigma^2 \Delta X}{\Delta\sigma^2} \quad (14)$$

I explain in section 5.3 how I then use my empirical estimates of dd_σ and dd_X plus a plausible range of values for $\frac{\sigma^2 \Delta X}{\Delta\sigma^2}$ to obtain a range of values of X .

The second method for empirically approximating X begins with the equation for dd_σ , combining equations 6-9 in the text. This is shown in equation 15, where recall $\Delta\sigma^2 = \sigma_L^2 - \sigma_H^2$.

$$\begin{aligned} dd_\sigma &= (Var(\Delta w)_{stay}^L - Var(\Delta w)_{move}^L) - (Var(\Delta w)_{stay}^H - Var(\Delta w)_{move}^H) \\ &= \frac{1}{4} \left(X \Delta\sigma^2 + \frac{(X\sigma_H^2)^2}{\left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma_H^2}\right)^2} - \frac{(X\sigma_L^2)^2}{\left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma_L^2}\right)^2} \right) \end{aligned} \quad (15)$$

This expression has 4 unknowns, X , $\bar{\theta}$, σ_L^2 and $\Delta\sigma^2$, and the goal is to empirically estimate X . I now take advantage of other aspects of the model, which can be easily estimated. Specifically, I turn to the probability that a group of workers stays with the incumbent. It is easy to write down the closed-form solution for this probability as a function of X , $\bar{\theta}$ and the relevant σ^2 . Furthermore, the probability of staying is easy to estimate empirically, although I note that I have not made the careful identification argument I made for dd_σ and dd_r in section 3. These exercise should therefore be taken with a large grain of salt.

I next derive an analytical solution for the probability that a worker in the

high variance ability group stays with the incumbent.

$$\begin{aligned}
\Pr(stay) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Pr(stay|\mu, s) d\mu ds = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\mu - v^*(s) + \bar{\theta}}{4\bar{\theta}} d\mu ds \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\mu - mX - s(1-X) + \frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma_H^2}}{4\bar{\theta}} d\mu ds \\
&= \frac{1}{4\bar{\theta}} \left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma_H^2} \right)
\end{aligned}$$

Similarly, we can derive the probability that someone in the low variance ability group stays. Furthermore, let γ_H and γ_L be the empirical probabilities that a worker stays with the incumbent firm for the high- and low-variance ability groups, respectively. We therefore have equations 16 and 17.

$$\gamma_H = \frac{1}{4\bar{\theta}} \left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma_H^2} \right) \quad (16)$$

$$\gamma_L = \frac{1}{4\bar{\theta}} \left(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma_L^2} \right) \quad (17)$$

First, these equations help in simplifying the expression for dd_σ (equation 15). We can plug in $4\bar{\theta}\gamma_H$ for $(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma_H^2})$ and $4\bar{\theta}\gamma_L$ for $(\frac{5}{2}\bar{\theta} - \sqrt{\frac{9}{4}\bar{\theta}^2 - X\sigma_L^2})$. In addition, I can use either of these two equations to solve for $\bar{\theta}^2$ and plug that back into equation 15. For example, using equation 17, $\bar{\theta}^2 = \frac{-X\sigma_L^2}{16\gamma_L^2 + 4 - 20\gamma_L}$. Then, I can use the two equations to solve for $\frac{\sigma_H^2}{\sigma_L^2} = \frac{16\gamma_H^2 + 4 - 20\gamma_H}{16\gamma_L^2 + 4 - 20\gamma_L}$. For ease of notation, let us denote this ratio with p . What I am doing here is using the differences in mobility probabilities to infer something about the underlying variances in ability, assuming that $\bar{\theta}$ and X are held constant across the two groups. Finally, if $\Delta\sigma^2 = \sigma_L^2 - \sigma_H^2$ and $p = \frac{\sigma_H^2}{\sigma_L^2}$, we must have $\sigma_L^2 = \frac{\Delta\sigma^2}{1-p}$.

Making these substitutions we get expression 18 where X is a function of empirical knowns and $\Delta\sigma^2$. Estimating $\Delta\sigma^2$ requires an additional empirical assumption, which I explain in the text.² I also discuss how a specifically empirically estimate dd_σ , γ_L , γ_H , and $\Delta\sigma^2$, which are all I need to find the

²I also note that I could perform an analogous exercise using dd_X . However, this would require an assumption about $X^A - X^S$. At present, I do not believe I have the ability to estimate that expression.

solution to X .

$$4dd_\sigma = X\Delta\sigma^2\left(1 - \frac{1}{16(1-p)}\left(\frac{p^2}{\gamma_H^2} - \frac{1}{\gamma_L^2}\right)(16\gamma_L^2 + 4 - 20\gamma_L)\right) \quad (18)$$

Figure 2: AFQT Distribution and Mean, by Job Entry Cohort

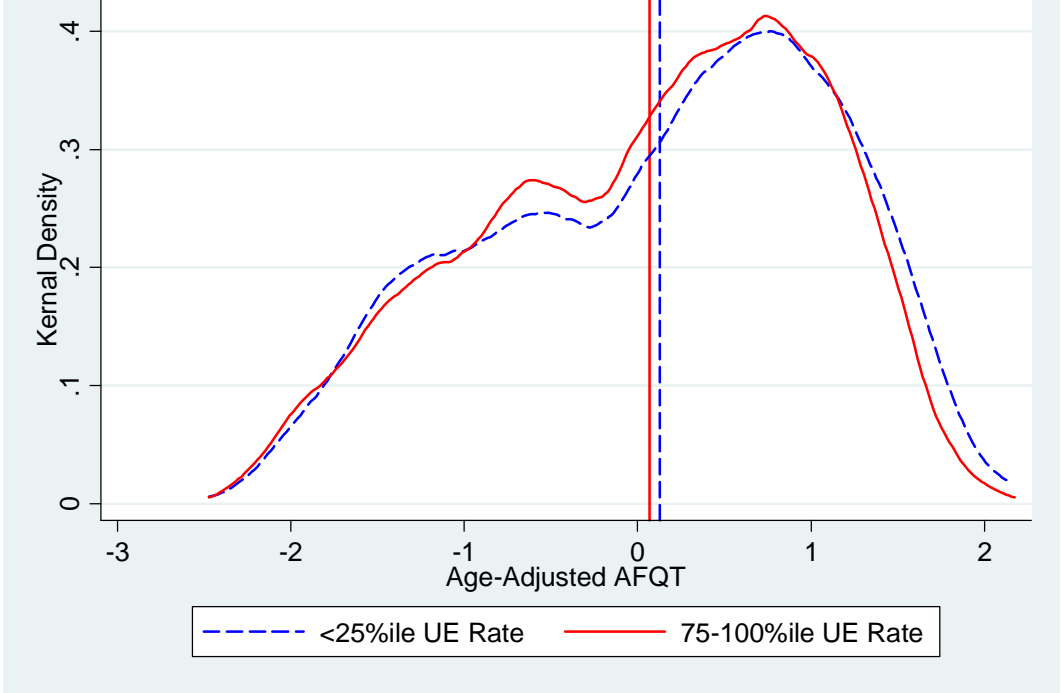


Table 1: Characteristics by National Job-Entry Cohort¹

	Highest Grade	AFQT	First Wage Residual ²
A: Means (relative to lowest quartile)			
75-100%ile UE Rate	-0.397** [0.0995]	-0.0573+ [0.0317]	-0.0346+ [0.0180]
50-75 %ile UE Rate	-0.169 [0.106]	-0.0562+ [0.0291]	-0.00602 [0.0140]
25-50%ile UE Rate	-0.345* [0.160]	-0.0905* [0.0340]	0.0242 [0.0177]
B: Squared Residuals ³ (relative to lowest quartile)			
75-100%ile UE Rate	-1.684** [0.287]	-0.0739+ [0.0379]	-0.0679** [0.0174]
50-75 %ile UE Rate	-1.169** [0.217]	-0.0382 [0.0334]	-0.0479* [0.0176]
25-50%ile UE Rate	-1.341** [0.408]	-0.0481 [0.0375]	-0.0153 [0.0237]
Observations	13,312	12,852	13,312

Robust standard errors in brackets clustered by job-entry year

+ significant at 10%; * significant at 5%; ** significant at 1%

Notes: Sample includes the cross-section, non-military male sample of the NLSY from 1979-2000. All jobs with at least one wage observation meeting the following criteria are included: post-transition, full-time, non-self-employed, non-enrolled work.

1. Job-Entry Cohort is defined as the national unemployment rate in the year the worker started his job. The cutoffs for the quartiles are the following: <5.5, 5.5-6.1, 6.2-7.2, >7.2.

2. The first log wage at the employer is residualized on the following variables: education dummies, geographic region and urban status dummies, marital status, a quadratic in age, black and hispanic dummies.

3. Characteristics are regressed on unemployment rate fixed effects and residuals are obtained.

Table 2: Characteristics of Movers by Communication Group¹

	AFQT		Highest Grade Completed		Previous Tenure		Log Wage Change		Pr(Layoff/Fired cond'l on move) ²
Move	-0.120**	-0.0872**	-0.178**	-0.0551	-0.640**	-0.776**	0.0502**	0.0143	--
	[0.0230]	[0.0288]	[0.0485]	[0.0464]	[0.0506]	[0.0807]	[0.00862]	[0.0146]	
Communication Group (rel to lowest quartile):									
75-100%ile comm grp	0.0256		0.14		0.360**		0.0668**		-0.179**
	[0.0396]		[0.109]		[0.136]		[0.0198]		[0.0177]
50-75 %ile comm grp	-0.0613		-0.296*		0.286*		0.0785**		-0.110**
	[0.0513]		[0.120]		[0.127]		[0.0224]		[0.0163]
25-50%ile comm grp	0.0389		0.0687		-0.0147		0.0101		-0.0709**
	[0.0617]		[0.0935]		[0.154]		[0.0238]		[0.0260]
Observations	30,346	30,346	31,435	31,435	31,435	31,435	28,974	28,974	7,988
R-Squared	0.004	0.092	0.001	0.145	0.008	0.013	0.004	0.008	0.024

Robust standard errors in parentheses, clustered by occupation

+ significant at 10%; * significant at 5%; ** significant at 1%

Notes: Sample includes the cross-section, non-military male sample of the NLSY from 1979-2000. All jobs with at least one wage observation meeting the following criteria are included: post-transition, full-time, non-self-employed, non-enrolled work.

1. Communication group is the quartile importance of "Communicating with Persons Outside Organization" from O*Net. Regressions include the main effect of quartile communication group as well as the interactions with moving shown above.
2. Regression only includes workers who moved jobs in between wage observations.

**Table 3: Spread of Log Wage Change Residuals¹ by Job-Entry UE Rate and Communication Group
(Dependent Variable: Squared Residuals²)**

	Full Sample		College+		< College	
	All	Low Ten ³	All	Low Ten ³	All	Low Ten ³
Stay at Employer	-0.0636** [0.0168]	-0.0398+ [0.0206]	-0.0865* [0.0339]	-0.0342 [0.0474]	-0.0497* [0.0189]	-0.0308 [0.0206]
Unemployment rate ⁴ (relative to lowest quartile):						
Stay*75-100%ile UE Rate	-0.0417** [0.0112]	-0.0562** [0.0142]	-0.0119 [0.0407]	-0.0475 [0.0417]	-0.0513** [0.0129]	-0.0583** [0.0129]
Stay*50-75%ile UE Rate	-0.0107 [0.0127]	-0.0344+ [0.0169]	0.0456 [0.0322]	-0.00845 [0.0396]	-0.0262+ [0.0153]	-0.0418* [0.0166]
Stay*25-50%ile UE Rate	-0.0184 [0.0195]	0.00018 [0.0234]	0.0840+ [0.0437]	0.0961+ [0.0494]	-0.0400+ [0.0200]	-0.0207 [0.0270]
Communication Group ⁵ (relative to lowest quartile):						
Stay*75-100%ile comm grp	-0.03 [0.0223]	-0.0833* [0.0367]	-0.0453 [0.0318]	-0.134* [0.0596]	-0.0376 [0.0301]	-0.0666 [0.0520]
Stay*50-75 %ile comm grp	-0.0226 [0.0309]	-0.0355 [0.0351]	0.00643 [0.0289]	-0.0774 [0.0467]	-0.0317 [0.0338]	-0.0304 [0.0370]
Stay*25-50%ile comm grp	0.00589 [0.0193]	0.00967 [0.0217]	-0.00269 [0.0980]	0.034 [0.0474]	0.00369 [0.0229]	0.00328 [0.0219]
Observations	28,416	15,240	4,906	2,414	23,510	12,826
R-squared	0.014	0.015	0.013	0.019	0.015	0.014

Standard errors clustered by job entry year.

+ significant at 10%; * significant at 5%; ** significant at 1%

Notes: Sample includes the cross-section, non-military male sample of the NLSY from 1979-2000. All jobs with at least one wage observation meeting the following criteria are included: post-transition, full-time, non-self-employed, non-enrolled work.

1. Log wage changes are regressed on the following and squared residuals are obtained: Job-entry quartile dummies, communication group quartile, an indicator for staying and interactions with each of the previous, a quadratic in age, dummy variables for year tenure, dummy variables for next-period tenure, change in squared tenure, final tenure with previous employer if moves (otherwise 0), education dummies, geographic region and urban status dummies, marital status, actual labor market experience, year effects, black and hispanic dummies, current industry and occupation dummies, and dummies for having changed industry or occupation, an indicator for having always been post-transition in the sample period, cumulative training, and interactions of communication group in t-1 and t.

2. Regressions also include controls for race and education, the main effects of communication group and job-entry group quartile, and the interactions of communication group in t-1 and t.

3. Restricted to observations with tenure less than 2 years (in first wage observation).

4. Job-Entry Cohort is defined as the national unemployment rate in the year the worker started his job. The cutoffs for the quartiles are the following: <5.5, 5.5-6.1, 6.2-7.2, >7.2.

5. Communication group is the quartile importance of "Communicating with Persons Outside Organization" from O*Net.

Table 4: Probability of Staying by Job-Entry UE Rate and Communication

	Group ¹		Low Tenure ²	
	Full Sample		1	2
	1	2	1	2
Unemployment rate ³ (relative to lowest quartile):				
75-100%ile UE Rate	-0.0481 [0.0374]	-0.0419 [0.0381]	-0.0246 [0.0386]	-0.0285 [0.0396]
50-75%ile UE Rate	-0.0501 [0.0354]	-0.049 [0.0357]	-0.0372 [0.0339]	-0.0399 [0.0343]
25-50%ile UE Rate	-0.0269 [0.0347]	-0.0265 [0.0348]	0.00934 [0.0216]	0.00176 [0.0217]
Communication Group ⁴ (relative to lowest quartile):				
75-100%ile comm grp	0.0167 [0.0129]	0.0114 [0.0126]	0.0161 [0.0135]	0.0107 [0.0139]
50-75 %ile comm grp	0.0038 [0.0135]	0.00767 [0.0131]	0.00691 [0.0141]	0.0105 [0.0139]
25-50%ile comm grp	0.0126 [0.00800]	0.0163+ [0.00814]	0.0253* [0.00981]	0.0292** [0.00884]
Log(wage)		0.187** [0.00788]		0.203** [0.0121]
Age-Adj AFQT		-0.00575+ [0.00323]		-0.00308 [0.00630]
Observations	28,974	27,979	15,599	15,089
R-squared	0.154	0.177	0.121	0.149

Standard errors clustered by job entry year.

+ significant at 10%; * significant at 5%; ** significant at 1%

Notes: Sample includes the cross-section, non-military male sample of the NLSY from 1979-2000. All jobs with at least one wage observation meeting the following criteria are included: post-transition, full-time, non-self-employed, non-enrolled work.

1. An indicator for staying with incumbent employer between periods is regressed on job-entry group and communication group dummies (reported here) as well as the following controls: a quadratic in age, dummy variables for year tenure, education dummies, geographic region and urban status dummies, marital status, actual labor market experience, year effects, black and hispanic dummies, current industry and occupation dummies, an indicator for having always been post-transition in the sample period, and cumulative training.

2. Restricted to observations with tenure less than 2 years (in first wage observation).

3. Job-Entry Cohort is defined as the national unemployment rate in the year the worker started his job. The cutoffs for the quartiles are the following: <5.5, 5.5-6.1, 6.2-7.2, >7.2.

4. Communication group is the quartile importance of "Communicating with Persons Outside Organization" from O*Net.

Table 5: Estimates of the Degree of Asymmetric Learning

	Method 1 ¹			Method 2 ²
	Value of $\Delta\sigma^2/\Delta\sigma$			
	0.8	1	1.2	
Degree of Asymmetry (X) ¹	0.540	0.674	0.809	0.688
Reduction of Expectation Error, relative to 100% by Incumbent:				
After 1 Year ³	0.460	0.326	0.191	0.312
After 5 Years ⁴	0.810	0.707	0.541	0.694

1. $X = (\beta_\sigma/\beta_X) * (\Delta X \sigma^2 / \Delta \sigma^2)$. β_σ and β_X are taken from the low tenure regression from the full sample in table 3 and are the coefficients on Stay*75-100%ile UE Rate and Stay*75-100%ile comm grp, respectively. See text for more detail.

2. X shown in equation 18 in the appendix. For β_σ , I use the coefficient on Stay*75-100%ile UE Rate from the low tenure regression on the full sample in table 3. For the probability that a worker in the high-variance ability group stays, I use the sample mean in the low tenure sample. For the probability in the low-variance ability group sample, I add the coefficient on 75-100%ile UE Rate from the low tenure regression in table 4 to the sample mean. For the difference in variances in ability, I use the difference in variance of AFQT in the low tenure sample (estimating using squared residuals). See text for more detail.

3. Corresponds to the speed of learning parameter in Lange (2007), K_1 . Here equals 1-X.

4. With the same noise structure on the learning parameters in each period, the expectation error is reduced on average by $5K_1/(1+4K_1)$ after 5 years.

Appendix Table B1: Data Description

Variable	Description	Codes
Wage	NLSY created measure of hourly rate of pay	CPI adjusted to 2000 dollars, missing if <\$1 or >\$500.
AFQT	Section2 + Section3 + Section4 + .5*Section5	For the entire NLSY sample, I create means and standard deviations by birth year then standardize each score by these.
Occupation	3-digit 1970 codes, recoded to 1 digit (manager, professional, sales, clerical, craft, operative, transport, labor, farm, service)	Occ/ind are constant w/in a job. Defined by last observation w/in a job. If missing, use second-to-last obs and so forth.
Industry	3-digit 1970 codes, recoded to the standard 15 categories	
Education	Followed responses to education questions year by year, then create constant measure within person equal to max	in regressions, categories are Prof, MA, BA/BS, AA, HS, GED and dropout, also use total years school "Highest Grade"
Tenure	NLSY created measure, weeks tenure as of int date	
Final Tenure	Constant within job	Tenure at last observation of job or number of weeks b/w start and stop date if tenure is missing (use int date as stop if currently working at last obs of job)
Start Date	Constant within job	Equals reported start date or, if missing, interview date at first observation
Stop Date	Constant within job	Reported stop date or, if missing, int date of last observation if not final job
Reason	Survey question, reason left job	Missing if stop year before 1984, 0 if first job, o.w. categories are: layoff/fired, plant closing, end temp/seasonal/program, or quit
ue rate group	National unemployment rate quartile in start year (<5.6, 5.6-6.2, 6.2-7.2, >7.2)	Missing if first job, equals ue rate for previous job if in first wage observation at job ("movers")
comm group		

Appendix Table B.2: Importance of "Communicating with Persons Outside Organization"¹

Occupation ²	Importance	Level
Sales Workers	4.52	4.24
Personnel and labor relations workers	4.50	3.50
Vocational/educational counselors	4.50	4.50
Receptionists	4.33	3.66
Social and recreational workers	4.28	4.30
Teachers (exc post-sec)	4.18	3.90
Computer specialists	4.00	4.50
Lawyers and judges	4.00	5.03
Research workers	3.98	4.40
Librarians/archivists/curators	3.96	4.13
Nurses/dietitians/therapists	3.94	3.89
Management and Admin	3.94	4.23
Secretaries	3.83	2.96
Accountants/Architects	3.71	4.75
Religious workers	3.66	4.66
Real estate appraisers	3.66	4.16
Food service workers	3.62	2.74
Other Clerks (exc secretaries)	3.60	2.67
Physicians, dentists, and related	3.59	4.11
Social Scientists	3.50	4.30
Clerical workers	3.49	2.81
Health service workers	3.35	2.97
Protective service workers	3.35	2.69
Farm/Forest/Home MGMT	3.31	4.33
Writers, artists, and entertainers	3.31	3.31
Health technologists and technicians	3.18	3.13
Personal service workers	3.15	2.45
Engineers	3.12	4.27
Post-secondary teachers	3.11	3.32
Farm Managers and Laborers	3.04	2.93
Operations Scientists	3.00	3.83
Life and Physical Scientists	2.87	3.90
Transport Equipment Operatives	2.81	2.11
Mathematical specialists	2.30	2.75
Mechanics and Repairmen	2.21	1.78
Technicians (exc health, eng., and sci.)	2.17	1.82
Office machine operators	1.96	1.51
Laborers (exc farm)	1.92	1.36
Private hh workers	1.90	1.61
Engineering and science technicians	1.68	1.27
Cleaning service workers	1.55	1.27
Craftsman	1.54	1.04
Operatives (exc transport)	1.44	0.83
Textile Operatives	1.26	0.66
Precision Machine Operatives	1.26	0.52

1. Taken from O*Net 4.0 activities, which have values for SOC occupation codes. The crosswalk between these codes and 1970 Census occupation coded is available from the author upon request.

2. Occupations are somewhat aggregated 3-digit 1970 Census Occupation codes. Aggregation is available from the author upon request.

Appendix Table B.3: Summary Statistics

	Job Entry Cohort ¹						Communication Group ²			
	Full Sample		75-100%ile UE		<25%ile UE		75-100%ile		<25%ile	
	(n=28,794)		Rate		Rate		Importance		Importance	
	mean	stdev	mean	stdev	mean	stdev	mean	stdev	mean	stdev
Log Wage Change	0.04	0.36	0.04	0.35	0.03	0.36	0.05	0.38	0.02	0.34
Entry UE Rate	6.82	1.28	8.34	1.02	5.16	0.32	6.79	1.28	6.88	1.28
Communication importance	2.64	1.18	2.58	1.18	2.74	1.17	4.11	0.26	1.20	0.14
Stay	0.65	0.48	0.65	0.48	0.59	0.49	0.71	0.45	0.62	0.49
Low Tenure	0.54	0.50	0.54	0.50	0.62	0.49	0.48	0.50	0.56	0.50
Age	28.65	5.23	26.64	4.80	33.23	4.65	29.58	5.06	27.99	5.23
Black	0.12	0.32	0.11	0.32	0.12	0.33	0.09	0.28	0.13	0.34
Hispanic	0.07	0.26	0.09	0.28	0.07	0.26	0.05	0.22	0.07	0.26
Age-Adj AFQT	0.14	0.97	0.16	0.94	0.16	1.00	0.58	0.85	-0.12	0.92
Tenure	2.95	3.26	3.27	3.64	2.04	2.00	3.33	3.45	2.79	3.15
Ten pd 2	3.28	3.54	3.57	3.94	2.46	2.26	3.74	3.72	3.09	3.43
Change in tenure squared	3.72	15.25	3.84	16.63	3.27	7.69	4.25	17.35	3.47	14.65
Final Tenure	33.83	85.59	32.75	86.87	34.95	73.14	34.08	93.66	33.55	82.03
Accumulated Experience	198.14	223.04	119.35	163.67	381.11	287.30	179.98	219.93	210.04	221.80
Highest degree:										
Dropout	0.13	0.34	0.13	0.34	0.13	0.34	0.04	0.20	0.18	0.38
GED	0.11	0.31	0.10	0.30	0.13	0.33	0.06	0.23	0.14	0.35
HS	0.54	0.50	0.56	0.50	0.49	0.50	0.47	0.50	0.61	0.49
AA	0.05	0.22	0.05	0.22	0.04	0.20	0.06	0.25	0.04	0.20
BA/BS	0.14	0.35	0.14	0.34	0.16	0.37	0.30	0.46	0.03	0.16
MA	0.02	0.14	0.02	0.13	0.03	0.18	0.05	0.21	0.00	0.06
Prof	0.01	0.10	0.01	0.09	0.01	0.12	0.02	0.15	0.00	0.01
Geographic region:										
Northeast	0.18	0.38	0.18	0.38	0.15	0.36	0.21	0.41	0.16	0.36
North Central	0.28	0.45	0.26	0.44	0.30	0.46	0.28	0.45	0.29	0.45
South	0.36	0.48	0.37	0.48	0.36	0.48	0.34	0.47	0.37	0.48
West	0.18	0.39	0.19	0.39	0.18	0.39	0.17	0.38	0.18	0.39
Missing region	0.00	0.05	0.00	0.05	0.00	0.06	0.00	0.06	0.00	0.04
Urbanicity										
not SMSA	0.24	0.43	0.25	0.43	0.20	0.40	0.17	0.38	0.30	0.46
SMSA, not central city	0.35	0.48	0.34	0.47	0.42	0.49	0.36	0.48	0.34	0.47
SMSA, central city unknown	0.24	0.43	0.25	0.44	0.20	0.40	0.27	0.44	0.22	0.42
SMSA, central city	0.13	0.34	0.12	0.32	0.15	0.36	0.15	0.36	0.11	0.31
Missing urbanicity	0.04	0.19	0.04	0.21	0.03	0.16	0.04	0.21	0.03	0.17
not married	0.53	0.50	0.57	0.50	0.46	0.50	0.50	0.50	0.53	0.50
married	0.47	0.50	0.43	0.50	0.54	0.50	0.50	0.50	0.47	0.50
Missing marital	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cumulative training	1.13	1.48	0.94	1.33	1.62	1.73	1.44	1.68	0.87	1.25
Changed industry	0.22	0.42	0.23	0.42	0.26	0.44	0.17	0.38	0.25	0.43
Changed occupation	0.21	0.41	0.22	0.42	0.23	0.42	0.18	0.38	0.24	0.43
Always post-transition	0.24	0.43	0.23	0.42	0.16	0.37	0.22	0.41	0.28	0.45

Notes: Sample includes the cross-section, non-military male sample of the NLSY from 1979-2000. All jobs with at least one wage observation meeting the following criteria are included: post-transition, full-time, non-self-employed, non-enrolled work.

1. Job-Entry Cohort is defined as the national unemployment rate in the year the worker started his job. The cutoffs for the quartiles are the following: <5.5, 5.5-6.1, 6.2-7.2, >7.2.

2. Communication group is the quartile importance of "Communicating with Persons Outside Organization" from O*Net.