Training Contracts, Worker Overconfidence, and the Provision of Firm-Sponsored General Training

Author: Mitchell Hoffman – Berkeley
Presenter: Matthew Chao – Caltech

February 21, 2012

Introduction

- "Hold-up" problem in training: If workers can't credibly commit to stay with a firm after being trained, firms will under-invest in training.
- A question of optimal human capital policy.
- Becker 1964: suggests workers pay for training themselves, but this may be impractical; in addition, empirically this is rare.
- Training contracts: firms pay for training, and workers agree to stay for a certain period. Penalties are incurred for violations.
- Types of workers: truckers, policemen, firefighters, nurses, pilots, securities brokers, federal employees, etc.

Research Questions

Key questions:

- How do training contracts affect firm training, worker selection, and worker turnover?
- How do training contracts impact firm profits?
- How do training contracts impact worker welfare?

Follow-up questions:

- How do worker beliefs about their future productivity modulate their decision to stay or quit under different training contracts?
- What is optimal firm policy given worker beliefs (e.g. overconfidence)?

Data Source

Data from a leading trucking firm in the U.S.
Why trucking?

1. Drivers are paid piece-meal per-mile-driven (productivity is easily measurable).
2. Variation in productivity across workers that is unknown ex-ante.
3. High turnover resulting in good identification on retention analysis.
4. Standard training: obtaining a commercial driver's license – classroom, simulator, actual driving. Courses are often required and last 2-4 weeks, including 148 hours of instruction with at least 44 hours of actual driving time.
5. In an informal survey, many such firms use training contracts.

Dataset Overview

His particular dataset (one firm):

- Pre-treatment: free training with no contractual obligations.
- Early 2000s: a newly promoted manager suggested training contracts: stay 12 months or pay a penalty. [Exogenous?]
- Contract phased into different training schools at different times based on state-by-state approvals. [Exogenous?]
- Approx. 5 years later: 18 month contract with higher quitting penalty that decreases with tenure.
- For one subset, also includes weekly panel data on subjective productivity forecasts (drivers are paid mostly piece rate per-mile-driven).

Structural Assumptions

Dynamic model of turnover and belief formation:

- Worker productivity is unknown and learned over time (following Jovanovic 1979).
- Workers form expectations of future productivity and earnings to decide whether to quit.
- Workers may have biased priors and update faster/slower than a Bayesian.

The model can:

- Replicate the quit-tenure curve, productivity-curve, and belief-tenure curve; learning and overconfidence are key.
- Run counterfactual simulations to show changes in worker welfare and firm profits through the training contracts.
- Run simulations on the effect of changes in overconfidence on firm profits, worker welfare, and retention rates.

Preview of Results / Main Contributions

- Training contracts reduce worker turnover.
- Long-term field evidence of absolute overconfidence that increases firm profits and decreases worker welfare.

Literature Review

- Firm-sponsored training (Acemoglu and Pischke 1999), e.g. tuition reimbursement (Cappelli 2004; Lynch and Black 1998; Manchester 2009, 2011).
- Overconfidence (Moore and Healy 2008, De Bondt and Thaler 1995, Larkin and Leider 2011, etc.)
- Learning models, e.g. wage growth (Harris and Holmstrom 1982), wage discrimination (Altonji and Pierret 2001), worker learning (Bojilov 2011), etc.
- Using subjective beliefs to evaluate decisions, e.g. Manski (2000), Wang (2010), van der Klaauw and Wolpin (2008).

Data 1

Treat contract implementation at different training schools as exogenous changes to worker training arrangements.

- "According to the Director of Driver Training, management had not been previously aware of the possibility of using a training contract."
- Training contract template requires state certification, which required different times per state.
- Staggered implementation: some as early as April 2002, others late 2002, and some never were approved.
- Penalty: $3500-$4000, regardless of when during the 12-month period the driver quit (or was fired).
- No changes to wage during contract implementations.

Data 2

- After several years, switched to 18-month contracts in order to increase retention of new drivers. Penalties were higher initially ($5000) but decreased with tenure ($62.50 per week).
- Each week, $12.50 of the $62.50 was removed from the worker paycheck, and repaid after the 78 weeks (18 months) as a bonus.
- No changes to wage during contract implementations.
- No initial deposit; payment made upon early exit (and potentially referred to collection agencies).

Data 3

Table 1: Summary Statistics

Panel A: All New Drivers at Firm A

Variable	No Contract	12-Month Contract	18-month Contract
African-American	0.19	0.19	0.19
Hispanic	0.07	0.04	0.03
Female	0.08	0.08	0.08
Married	0.35	0.35	0.38
Age	37.27	37.00	37.25
Online Application	0.47	0.55	0.70
Smoker	0.30	0.42	0.40
Drivers	$0.09N$	$0.73N$	$0.18N$

Panel B: Drivers in Data Subset

Variable	Obs	Mean	Min	Max
African-American	895	0.11	0	1
Hispanic	895	0.02	0	1
Female	895	0.10	0	1
Married	895	0.41	0	1
Age	894	36.46	21.06	69.21
Number of Kids	895	0.96	0	7
Online Application	889	0.67	0	1
Smoker	787	0.46	0	1
Years of Schooling	895	12.85	9	18
High School Dropout	895	0.04	0	1
High School Graduate	895	0.40	0	1
Some College	895	0.34	0	1
Technical School	895	0.14	0	1
College Degree or More	895	0.08	0	1
Credit Score	784	585.96	407	813
No Credit Score	895	0.12	0	1

Reduced Form Analysis 1

Estimate the *Cox Proportional Hazard Rate* of quitting:

$$\log(h_{it\tau cs}) = \alpha_t + \beta_1 * SCONTRACT_{sc} + \beta_2 * LCONTRACT_{sc} + \beta_3 * UNEMP_{s\tau} + \beta_4 \bar{y_{it}} + \gamma_\tau + \delta_c + \theta_s + X_i\lambda + \epsilon_{it\tau cs}$$

where:

- $h_{it\tau cs}$ is the quit hazard rate of driver i with t weeks of tenure in year τ that is part of cohort c (year of hire) attending training school s.
- $UNEMP_{s\tau}$ is unemployment rate in state s at time τ
- $\bar{y_{it}}$ is average productivity to date
- α_t is fixed effect for tenure t
- γ_τ is time FE. δ_c is cohort c FE, and θ_s is school FE.
- X_i are individual covariates and $\epsilon_{it\tau cs}$ is error
- β_1 and β_2 represent effects of the two contract types on quit hazard.

Reduced Form Analysis 2

Table 2: Impact of the Training Contracts on Quitting – Cox Model, Diff-in-Diff

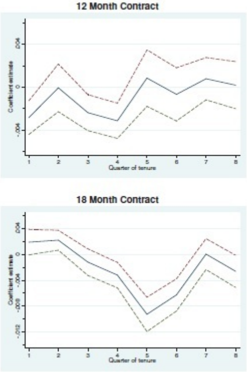
	(1)	(2)	(3)	(4)	(5)	(6)
12m contract	-0.179***	-0.173***	-0.208***	-0.202***	-0.182***	
	(0.046)	(0.046)	(0.052)	(0.055)	(0.058)	
18m contract	-0.108*	-0.118*	-0.116*	-0.127*	-0.107	
	(0.062)	(0.062)	(0.069)	(0.071)	(0.074)	
State unemployment rate		-0.055***	-0.048***	-0.062***	-0.049***	-0.054***
		(0.011)	(0.013)	(0.013)	(0.014)	(0.014)
Avg miles to date				-0.060***	-0.048***	-0.049***
				(0.003)	(0.004)	(0.004)
12m contract * ($wks \leq 52$)						-0.371***
						(0.074)
12m contract * ($52 < wks \leq 78$)						0.066
						(0.081)
12m contract * ($wks > 78$)						0.043
						(0.090)
18m contract * ($wks \leq 52$)						-0.021
						(0.087)
18m contract * ($52 < wks \leq 78$)						-0.891***
						(0.112)
18m contract * ($wks > 78$)						-0.100
						(0.132)
Time FE (yr)	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE (yr of hire)	Yes	Yes	Yes	Yes	Yes	Yes
Training School FE	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	No	No	No	Yes	Yes
Observations	M	M	$0.89M$	$0.89M$	$0.79M$	$0.79M$

Notes: An observation is a driver-week. The regressions are Cox proportional hazard models, with standard errors clustered by school-week of hire in parentheses. Driver tenure is controlled for non-parametrically. State unemployment is a state's annual unemployment rate. Demographic controls include gender, race dummies, marital status, and driver age. Average miles to date is a driver's average weekly productivity to date and is given in terms of hundreds of miles driven per week. Column (3) differs from column (2) differs in that it restricts the sample to be the same as in column (4). The exact M is withheld to protect firm confidentiality, $M >> 100,000$. * significant at 10%; ** significant at 5%; *** significant at 1%

Reduced Form Analysis 3

OLS estimates using quarter-tenure blocks yields similar estimates:

1. Relative to having no contract, 12-month contracts yield lower quitting in the 4th quarter but higher quitting in the 5th quarter
2. 18-month contracts yield lower quitting in quarters 4-6 but higher quitting afterward quarter 6.



Alternative Explanations

Endogeneity concerns:

- Worker sorting into schools in states without training.
- Varied contract enforcement (e.g. by tenure or performance).
- State-by-state implementation was endogenous to factors that affect quitting (e.g. unemployment rates, etc.).

Testing Incentive Effectives versus Selection:

- Rates of firing don't change in response to contracts.
- Neither do productivity characteristics of workers, though selection may have occurred on several other characteristics (smokes, applied online, Hispanic).
- If there is selection on unobservables, the coefficient magnitudes should change when controlling for productivity (in particular, they should decrease if contracts cause positive selection in favor of "good" drivers)

Worker Beliefs

Key Question: Can incorporating worker beliefs about their future productivity better explain their behavior under different training contracts?
Assumptions:

- Workers are initially uncertain about their productivity.
- Workers gain information through weekly productivity signals.
- Workers that learn they are less productive are more likely to quit.

Corroborating evidence:

1. At every point in time, less productive workers are more likely to quit. Quitting drivers receive lower average earnings in prior weeks than non-quitting drivers.
2. With controls, past average productivity reduces the hazard of quitting.

Survey

Procedure:

1. Each week, asked over Qualcomm message system (in their truck): "About how many paid miles do you expect to run during your next pay week?"
2. Drivers typed responses. Told they weren't shared with the company [concern?]
3. $5 each week for completing the survey; no other incentives.
4. Average response rate across all drivers and weeks: 21 percent.
5. 699 drivers, of which 61% responded at least once.

Survey

Respondent statistics:

- Survey incentives for accurate guessing didn't decrease overconfidence in a survey at a different firm.
- Women and minorities were less likely to respond.
- Older drivers and those with higher average productivity were more likely to respond; within-driver, a more productive week increased the likelihood of responding.

Beliefs and Actual Productivity

Estimate whether productivity beliefs predict productivity:

$$y_{i,t} = \alpha + \beta b_{i,t-1} + \gamma \bar{y}_{i,t-1} + X_i \delta + \epsilon_{i,t}$$

where:

- $y_{i,t}$ is driver i's productivity in week t of his tenure with the company
- $b_{i,t-1}$ is previous week's productivity belief
- $\bar{y}_{i,t-1}$ is lagged average productivity to date
- X_i are controls, and $\epsilon_{i,t}$ are errors

Beliefs and Actual Productivity

Table 4: Do Productivity Beliefs Predict Productivity? OLS Regressions

	(1)	(2)	(3)	(4)	(5)
L. Pred miles	0.195***	0.066***	0.064***	0.079***	0.086***
	(0.023)	(0.016)	(0.016)	(0.022)	(0.023)
L. Avg miles to date		0.789***	0.689***		-0.188*
		(0.037)	(0.038)		(0.108)
Tenure FE	No	Yes	Yes	Yes	Yes
Demographic Controls	No	No	Yes	No	No
Education Controls	No	No	Yes	No	No
Work Type Controls	No	No	Yes	No	No
Subject FE	No	No	No	Yes	Yes
Observations	8,449	8,435	8,435	8,449	8,435
R-squared	0.05	0.17	0.18	0.29	0.29

Notes: The dependent variable is miles driven per week. An observation is a driver-week. Standard errors clustered by driver in parentheses. Demographic controls include gender, race dummies, marital status, and age bin dummies for the different age groups: 25-30, 30-35, 35-40, 40-45, 45-50, 50-55, 55-60, and 60-80. Education controls are dummies for high school graduate, some college, and college. Work type controls are dummies for different work configurations and for receiving any salary or activity-based pay. Productivity is given in terms of hundreds of miles driven per week. All drivers are from the same training school and were hired in late 2005 or 2006. * significant at 10%; ** significant at 5%; *** significant at 1%

Beliefs and Quit Rates

Estimate whether productivity beliefs predict quitting:

$$\log h_{i,t} = \alpha + \beta b_{i,t-1} + \gamma \bar{y}_{i,t-1} + X_i \delta + \epsilon_{i,t}$$

where $h_{i,t}$ is the quit hazard rate.

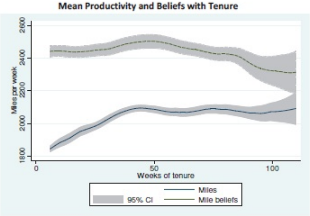
Beliefs and Actual Productivity

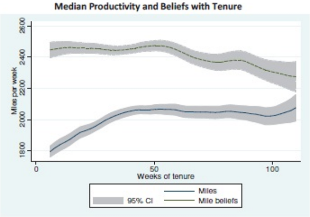
Table 5: Do Productivity Beliefs Predict Quitting?

	(1)	(2)	(3)	(4)	(5)
Predicted miles	-0.059***			-0.059***	-0.067***
	(0.018)			(0.020)	(0.021)
Avg miles to date		-0.079***	-0.112***	-0.002	-0.062
		(0.010)	(0.039)	(0.036)	(0.042)
Demographic Controls	No	Yes	Yes	No	Yes
Education Controls	No	Yes	Yes	No	Yes
Work Type Controls	No	Yes	Yes	No	Yes
Observations	8,500	38,381	8,500	8,500	8,500

Notes: An observation is a driver-week. The regressions are Cox proportional hazard models, where the dependent variable is quitting. Events where the driver is fired are treated as censored. Standard errors clustered by worker are in parentheses. Demographic controls include gender, race dummies, marital status, and age bin dummies for the different age groups: 25-30, 30-35, 35-40, 40-45, 45-50, 50-55, 55-60, and 60-80. Education controls are dummies for high school graduate, some college, and college. Work type controls are dummies for different work configurations and for receiving any salary or activity-based pay. Productivity is given in terms of hundreds of miles driven per week. All drivers are from the same training school and were hired in late 2005 or 2006. * significant at 10%; ** significant at 5%; *** significant at 1%

Characteristics of Worker Beliefs

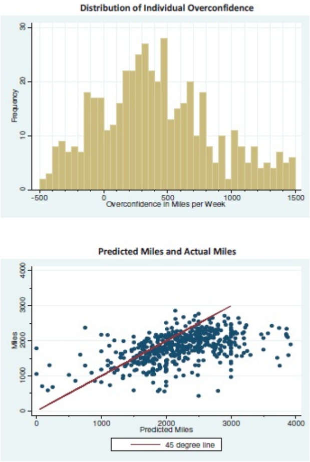




Notes: This figure analyzes actual and believed productivity for 699 drivers in the data subset. The figures are plotted using a local polynomial regression with an Epanechnikov kernel. A bandwidth of 5 weeks is used for the productivity data and a bandwidth of 7 weeks is used for the belief data. In the top figure, the productivity and belief data are collapsed into weekly means before local polynomial smoothing. In the bottom figure, the productivity and belief data are collapsed into weekly medians before local polynomial smoothing.

Characteristics of Worker Beliefs

Figure 7: Distribution of Overconfidence Across Drivers



Notes: This figure analyzes overconfidence among 699 drivers in the data subset. It presents reduced-form evidence on the distribution of overconfidence across drivers. The top figure plots a histogram of driver-level overconfidence, where overconfidence is defined as the difference between beliefs and productivity. A driver's overconfidence level is calculated by averaging over all the weeks with productivity beliefs and actual productivity. In the bottom figure, each driver is represented by a dot located at their average productivity and beliefs (averaged over all weeks the driver is observed).

Structural Model Assumptions

Initial assumptions:

- All previously discussed assumptions (e.g. piecerate pay, initially unknown productivity, etc.).
- Only decision is whether to quit (can't "unquit" in the future). No effort decision.
- Each week's miles driven is a noisy signal about true productivity.
- Worker is forward looking, but may have belief biases.
- Workers may overweight or underweight his prior when updating.
- Biases in priors diminish with new information but at different speeds.
- Heterogeneity in: productivity, beliefs, taste for the job, plus idiosyncratic shocks (e.g. a fight with a boss or co-worker).

Model Setup

- Baseline productivity $\eta \sim N(\eta_0, \sigma_0^2)$
- Paid piece-rate w_t that varies by tenure. Workers believe this profile remains constant over time.
- A worker's weekly miles are distributed $N(a(t) + \eta, \sigma_y^2)$.
- $a(t)$ is a learn-by-doing process (e.g. improving with experience).
- Weekly earnings $Y_t = w_t * y_t$.
- Outside option r_t (which may depend on the worker's tenure when he quits).
- Each period, decision d_t is made to stay $(d = 1)$ or quit.
- Workers in time t know past miles up to $t - 1$ but not in the current week.
- Workers are risk-neutral with discount factor δ.

Decision Model

Maximize expected utility:

$$V_t(\mathbf{x}_t) = \max_{d_t, d_{t+1}, \ldots} E_t \left(\sum_{s=t}^{\infty} \delta^{t-s} u_t (d_s, \mathbf{x}_s) \, | d_t, \mathbf{x}_t \right).$$

where x_t is the vector of state variables at time t, including past productivity, piece rate, training contract, taste, and confidence. Rewritten in Bellman form as:

$$V_t(\mathbf{x}_t) = \max_{d_t} E_t \left(u_t (d_t, \mathbf{x}_t) + \delta V_{t+1}(\mathbf{x}_{t+1}) | d_t, \mathbf{x}_t \right).$$

Decision Model

Per-period utility for staying or quitting are, respectively:

$$u_t(1, \mathbf{x}_t) = \alpha + w_t y_t + \varepsilon_t^S,$$

$$u_t(0, \mathbf{x}_t) = -k_t + \frac{r_t}{1-\delta} + \varepsilon_t^Q.$$

where:

- α represents the worker's taste for working in trucking
- k_t is a vector representing the penalty for quitting at tenure t
- ϵ_t^S and ϵ_t^Q are iid idiosyncratic unobserved error

Decision Model

Choice-specific value functions are thus:

$$
\begin{aligned}
V_t^Q &= -k_t + \frac{r_t}{1-\delta} + \epsilon_t^Q \equiv \overline{V}_t^Q + \epsilon_t^Q \\
V_t^S &= \alpha + E_t(w_t y_t | \mathrm{x}_t) + \delta E(V_{t+1}(\mathrm{x}_{t+1}) | \mathrm{x}_t) + \epsilon_t^S \equiv \overline{V}_t^S + \epsilon_t^S,
\end{aligned}
$$

The Bellman is thus:

$$
V_t(\mathrm{x}_t) = \max_{d_t \in \{0,1\}} \left(V_t^S(\mathrm{x}_t), V_t^Q(\mathrm{x}_t) \right).
$$

Belief Formation Model

Standard normal learning model:

- Worker beliefs in period t are a weighted sum of his priors and his average productivity signals.
- As t increases, weight shifts to the average productivity.

$$E(y_t|y_1,...,y_{t-1}) = \frac{\sigma_y^2}{(t-1)\,\sigma_0^2+\sigma_y^2}\eta_0 + \frac{(t-1)\,\sigma_0^2}{(t-1)\,\sigma_0^2+\sigma_y^2}\frac{\sum_{s=1}^{t-1} y_s - a(s)}{s-1} + a(t)$$

Augmented model:

- Allows overconfidence. Priors η_0 are drawn from $N(\eta_0+\eta_b,\sigma_0^2)$ not $N(\eta_0,\sigma_0^2)$
- Workers may perceive the noise of the signal as different from it's true value; e.g. signals have s.d. $\tilde{\sigma_y}$ instead of σ_y.
- $\eta_b > 0$ implies overconfidence. Speed at which weight on $(\eta_0+\eta_b)$ goes to 0 depends on $\tilde{\sigma_y}$.

$$E^b(y_t|y_1,...,y_{t-1}) = \frac{\widetilde{\sigma_y}^2}{(t-1)\,\sigma_0^2+\widetilde{\sigma_y}^2}(\eta_0+\eta_b) + \frac{(t-1)\,\sigma_0^2}{(t-1)\,\sigma_0^2+\widetilde{\sigma_y}^2}\frac{\sum_{s=1}^{t-1} y_s - a(s)}{s-1} + a(t)$$

Other Model Details

Timing for each period:

1. Workers form beliefs b_t given past earnings.
2. ϵ_t^S and ϵ_t^Q are realized. Worker makes decision.
3. If worker didn't quit, productivity y_t is realized.

$a(t)$ function represents increased productivity over time (a worker's weekly miles are distributed $N(a(t) + \eta, \sigma_y^2)$.

1. It's path is fully anticipated by workers
2. A function of tenure only

Reservation wage:

1. $r_t = r - \frac{6-t}{5} s_0$ for $t \leq 5$ and $r_t = r$ for $t > 5$.
2. Reflects that for the first 4-6 weeks, drivers receive a flat $375 per week while driving with an experienced colleague.
3. r is fixed using outside data and s_0, the skills gained during this time, are estimated.

Estimation

Use MLE. Likelihood function:

$$\begin{aligned} L_i &= \int L(d_{i1},...,d_{it},y_{i1},...,y_{it},b_{i1},...,b_{it}|\alpha,\eta_b) f(\alpha,\eta_b)\, d\alpha d\eta_b \\ &= \left(\prod_{s=1}^{t} L(y_{is}|y_{i1}...,y_{is-1})\right)\left[\int \prod_{s=1}^{t} L(d_{is}|y_{i1},...,y_{is-1},\alpha,\eta_b)\prod_{s=1}^{t} L(b_{is}|y_{i1}...,y_{is-1},\eta_b) f(\alpha,\eta_b)\, d\alpha d\eta_b\right] \\ &\equiv L_i^2 \int L_i^1(\alpha,\eta_b) L_i^3(\eta_b) f(\alpha,\eta_b)\, d\alpha d\eta_b \end{aligned}$$

where:

- $L_i^1(\alpha,\eta_b)$ is the contribution to likelihood from quitting decisions
- L_i^2 is the contribution from earnings realizations (past productivity)
- $L_i^3(\eta_b)$ is the contribution from subjective beliefs

Parameter identification:

- σ_0 reflects across individual differences; σ_y represents within individual differences. Identified off productivity data.
- s_0 (skill gain) is identified based on turnover during the first 5 weeks (assumes those that quit are low skill).
- Taste heterogeneity results from differences between individual quitting and model predictions.
- σ_b, s.d. of beliefs, comes from the difference in subjective expectations and mathematical expectations.
- $\tilde{\sigma_y}$ affects both subjective beliefs and quitting decisions. The faster one weighs past productivity, the smaller the $\tilde{\sigma_y}$; the faster initial overconfidence disappears, the smaller $\tilde{\sigma_y}$ will be.

Implementation

Numerical solutions:

- Discretize productivity into $K = 40$ values, in increments of 100 from $100 - 4000$ miles per week.
- Transitions between earnings state are:

$$\Pr(y_s^k | y_1, ..., y_{s-1}) = \Phi\left(\frac{y_s^k + .5 * kstep - E(y_s^k | y_1, ..., y_{s-1})}{\sqrt{\Omega_{s-1}}}\right) - \Phi\left(\frac{y_s^k - .5 * kstep - E(y_s^k | y_1, ..., y_{s-1})}{\sqrt{\Omega_{s-1}}}\right)$$

  where *kstep* is the distance between earnings realizations.

- For r, use the median full-time earnings from 2006 March Current Popular Survey of workers similar to the median driver; $r = \$640$ weekly.
- Assume weekly $\delta = 0.9957$ corresponding to an annual rate of 0.8.

Structural Results

Table 6: Baseline Structural Estimates

		No Bias	Belief Bias	Belief Bias 2 Types
τ	Scale param of idiosyncratic shock	1618	2206	3726
		(136)	(291)	(449)
	Productivity Parameters			
η_0	Mean of prior productivity dist	2464	2025	2024
		(9)	(17)	(18)
σ_0	Std dev of prior productivity dist	475	286	284
		(16.5)	(10.1)	(9.9)
σ_y	Std dev of productivity shocks	707	706	706
		(1.5)	(1.5)	(1.6)
	Skill Gain Parameter			
s_0	Value of skills gained in wks 1-5	14.9	8.6	31.7
		(3.5)	(5.9)	(9.0)
	Taste UH Parameters			
μ_1	Mass point 1 of taste UH	-248	-259	-736
		(9.0)	(12.6)	(41.9)
μ_2	Mass point 2 of taste UH	-106	-135	-150
		(14.6)	(12.1)	(10.0)
μ_3	Mass point 3 of taste UH	139	135	191
		(39.3)	(33.3)	(63.1)
p_1	Probability of type 1, taste	0.55	0.34	0.12
		(0.04)	(0.06)	(0.03)
p_2	Probability of type 2, taste	0.24	0.43	0.67
		(0.03)	(0.06)	(0.04)
	Belief Parameters			
σ_b	Std dev in belefs	299	298	271
		(0.3)	(0.3)	(0.3)
$\widetilde{\sigma_y}$	Believed std dev of productivity shocks	3650	1888	2068
		(134)	(82)	(81)
η_b	Belief bias		589	
			(22)	
$\eta_{1,b}$	Mass point 1 of belief UH			426
				(20)
$\eta_{2,b}$	Mass point 2 of belief UH			3649
				(41)
$p_{1,b}$	Probability of type 1, beliefs			0.94
				(0.16)
	Log-likelihood	-91064	-90865	-89882
	Number of workers	699	699	699

Notes: This table presents estimates of the structural parameters. The idiosyncratic shock, skill gain, and taste parameters are given in terms of dollars whereas the productivity and belief parameters are given in terms of miles. Standard errors are in parentheses and are calculated by inverting the Hessian. All specifications assume a normal learning model. A weekly discount factor of 0.9957 is assumed for workers and firms, corresponding to an annual discount factor of 0.8. The data are from 699 drivers in the data subset, all of whom face the 12-month training contract.

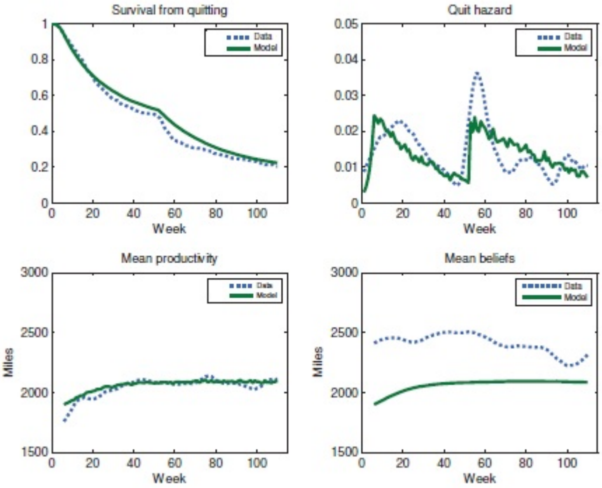
Structural Results

Table 7: Structural Estimates with Learning by Doing and Skill Accumulation

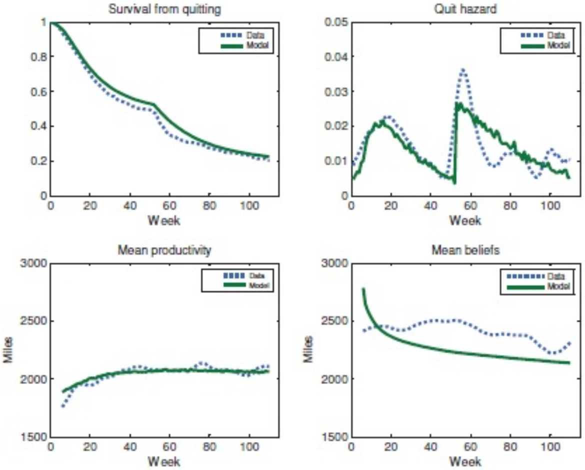
		No Bias	Belief Bias	Belief Bias 2 Types
τ	Scale param of idiosyncratic shock	1891	1605	4207
		(260)	(180)	(925)
	Productivity and Skill Parameters			
η_0	Mean of prior productivity dist	2307	1595	1742
		(28)	(41)	(33)
σ_0	Std dev of prior productivity dist	521	275	276
		(20.8)	(10.5)	(10.5)
σ_y	Std dev of productivity shocks	706	706	705
		(3.6)	(3.6)	(3.6)
b_1	Learning by doing level	251	485	343
		(33)	(38)	(34)
b_2	Learning by doing speed	0.08	0.11	0.07
		(0.01)	(0.02)	(0.01)
θ_1	Skill gain level	92	228	534
		(53)	(54)	(140)
θ_2	Skill gain speed	0.68	0.03	0.05
		(0.30)	(0.01)	(0.01)
	Taste UH Parameters			
μ_1	Mass point 1 of taste UH	-195	-345	-988
		(56.7)	(19.7)	(211.6)
μ_2	Mass point 2 of taste UH	-61	-81	27
		(58.1)	(30.0)	(58.5)
μ_3	Mass point 3 of taste UH	180	231	634
		(73)	(55)	(170)
p_1	Probability type 1	0.55	0.60	0.42
		(0.07)	(0.04)	(0.06)
p_2	Probability of type 2	0.25	0.19	0.35
		(0.06)	(0.04)	(0.06)
	Belief Parameters			
σ_b	Std dev in befefs	298	297	271
		(1.4)	(1.4)	(1.3)
$\widetilde{\sigma_y}$	Believed std dev of productivity shocks	3481	1295	1759
		(167)	(82)	(108)
η_b	Belief bias		722	
			(31)	
$\eta_{1,b}$	Mass Point 1 of belief UH			466
				(25)
$\eta_{2,b}$	Mass Point 2 of belief UH			3899
				(139)
p_1	Prob of Type 1			0.93
				(0.01)
	Log-likelihood	-91024	-90731	-89787
	Number of workers	699	699	699

Notes: This table presents estimates of the structural parameters. The idiosyncratic shock, skill gain, and taste parameters are given in terms of dollars whereas the productivity and belief parameters are given in terms of miles. Standard errors are in parentheses and are calculated by inverting the Hessian. All specifications assume a normal learning model. A weekly discount factor of 0.9957 is assumed for workers and firms, corresponding to an annual discount factor of 0.8. The data are from 699 drivers in the data subset, all of whom face the 12-month training

Structural Results



Structural Results



Simulations 1: Firm Profits and Worker Welfare

- Assess firm profits-per-worker and profits-per-truck under different training contracts.
- Simulate 3000 workers for up to 1300 weeks each. Both measures show profits greatest under 18-month contracts, and least under no contract.
- Also shows that, using experienced utility measures, worker welfare decreases under the contracts.

Table 8: Profits and Welfare Under Different Contracts

	No contract	12 month	18 month
Profits per worker	$363	$1,625	$1,875
Profits per truck	-$2,544	$1,856	$2,641
Welfare per worker	$159,580	$156,590	$156,375

Notes: This table presents profits per worker and profits per worker under the three different training contracts used by Firm A. Profits per worker and profits per firm are defined in Section 9 of the text. Profits are calculated assuming a Fixed Cost of $600 per week, a price of $1.80 per mile, a non-wage marginal cost of $1.16 per mile, a sunk cost of $2.50 per worker per week, and a marginal cost of training of $2,500, and a collection rate of 30%. Profits and welfare are calculated by simulating 40,000 workers under each of the three regimes. A weekly discount factor of 0.9957 is assumed for workers and firms, corresponding to an annual discount factor of 0.8. The model is estimated with 3 taste mass points and 2 overconfidence mass points. The model simulated has no learning by doing and assumes a flat outside option after the first 5 weeks.

Simulations 2: Debiasing Overconfidence

- Simulate the structural model but with overconfidence reduced (either by one half or completely).
- This reduces worker retention (more quit), because they see their future earnings for staying as lower.
- This reduces profits per worker and profits per truck.

Table 9: Counterfactual Simulations, No Contractual Response

	Baseline	50% debias	100% debias	Reveal ability
Retention at 20 wks	0.74	0.62	0.49	0.45
Retention at 40 wks	0.55	0.44	0.32	0.32
Retention at 60 wks	0.43	0.35	0.26	0.26
Welfare per worker	$156,590	$157,771	$158,708	$158,959
Profits per worker	$1,625	$793	-$83	$562
Profits per truck	$1,856	-$1,159	-$5,225	-$2,650
Ability at 20 wks	2,037	2,034	2,036	2,085
Ability at 40 wks	2,058	2,055	2,053	2,115
Ability at 60 wks	2,072	2,062	2,057	2,114

Notes: This table reports the reports the results of the counterfactual simulations described in the text, while assuming that training contracts are not adjusted in response. Under the 50% debias and 100% debias counterfactuals, worker overconfidence is reduced by 50% or 100% (by reducing the overconfidence mass points η_{b1}

Extras

- Optimize training contracts on a per-period basis, so that contracts aren't fixed over time.
- Banning training contracts as a restriction on the firm's optimization problem, with and without debiasing.

Takeaways

Key takeaways:

1. Plausibly exogenous contractual variation shows that contracts reduce quitting by 10-20 percent (equivalent to 2-4 percentage point increases in home state unemployment rates).
2. An incentive effect: little evidence of worker selection resulting from training contracts.
3. Subjective productivity beliefs correlate with actual quitting and productivity.
4. Workers are on average systematically overconfident, and persistently so in some respects.
5. A structural model of quitting demonstrates the role of overconfidence and learning in quitting decisions.
6. Without training contracts or overconfidence, firm profits from training drop substantially (but worker welfare increases).

Extensions

Extensions:

1. How does this vary empirically for high-skill positions?
2. Model competition amongst firms.
3. How else could worker overconfidence be relevant (e.g. what about offering not only piece-rates but also convex pay structures)?
4. Analyze optimal contract lengths and structures.