On the Role of Learning, Human Capital and Performance Incentives for Wages

Braz Camargo[1] Fabian Lange[2] Elena Pastorino[3]

[1]Sao Paulo-FGV [2]McGill University [3]Stanford University and Hoover Institution

ES Dynamic Structural Econometrics Conference 2024

Revisit Key Debate in Labor and Personnel Economics

- To align workers' incentives to firms' objectives, firms often link compensation to performance on the job
  - through bonuses, commissions, piece rates and similar
  - but for most, performance pay (PP) accounts for $<10\%$ of pay (not major component at any point over life cycle)
  - open question: do incentives for performance matter for the typical worker and if so why is PP so small?

Revisit Key Debate in Labor and Personnel Economics

- To align workers' incentives to firms' objectives, firms often link compensation to performance on the job
  - through bonuses, commissions, piece rates and similar
  - but for most, performance pay (PP) accounts for $<10\%$ of pay (not major component at any point over life cycle)
  - open question: do incentives for performance matter for the typical worker and if so why is PP so small?

- One reason: workers may already face strong *implicit* incentives for performance i.e. motivated to work hard
  - so as to convince potential employers, who are uncertain about their talent, their productivity is high

Revisit Key Debate in Labor and Personnel Economics

- To align workers' incentives to firms' objectives, firms often link compensation to performance on the job
  - through bonuses, commissions, piece rates and similar
  - but for most, performance pay (PP) accounts for $<10\%$ of pay (not major component at any point over life cycle)
  - open question: do incentives for performance matter for the typical worker and if so why is PP so small?

- One reason: workers may already face strong *implicit* incentives for performance i.e. motivated to work hard
  - so as to convince potential employers, who are uncertain about their talent, their productivity is high

- This well-known *career-concerns* (CC) argument (Holmström, 1982 and 1999) is common explanation
  - not only for why PP often makes up *only small portion of pay* but also for *how PP varies over time*

- Intuitively, as workers accumulate experience in labor mkt and their productivity becomes better known
  - implicit incentives from CC weaken so explicit incentives from PP should become more and more important

This Interpretation Raises a Puzzle

- Key prediction of CC models (Gibbons and Murphy, 1992): as ability is eventually learned
  - these *implicit* incentives for effort to "prove oneself" are progressively substituted by *explicit* incentives from PP
  - therefore PP as fraction of total pay *should increase with experience*

- This prediction has found some support from evidence on top-executive pay

This Interpretation Raises a Puzzle

- Key prediction of CC models (Gibbons and Murphy, 1992): as ability is eventually learned
  - these *implicit* incentives for effort to "prove oneself" are progressively substituted by *explicit* incentives from PP
  - therefore PP as fraction of total pay *should increase with experience*

- This prediction has found some support from evidence on top-executive pay but o.w. highly *counterfactual*

This Interpretation Raises a Puzzle

- Key prediction of CC models (Gibbons and Murphy, 1992): as ability is eventually learned
  - these *implicit* incentives for effort to "prove oneself" are progressively substituted by *explicit* incentives from PP
  - therefore PP as fraction of total pay *should increase with experience*

- This prediction has found some support from evidence on top-executive pay but o.w. highly *counterfactual*
  - PP as fraction of total pay tends to *decline over second half of life cycle* (precisely when it should be *increasing*)
  - in fact, profile of PP relative to total pay is *hump-shaped* with experience in a variety of public and firm-level data sets

This Interpretation Raises a Puzzle

- Key prediction of CC models (Gibbons and Murphy, 1992): as ability is eventually learned
  - these *implicit* incentives for effort to "prove oneself" are progressively substituted by *explicit* incentives from PP
  - therefore PP as fraction of total pay *should increase with experience*

- This prediction has found some support from evidence on top-executive pay but o.w. highly *counterfactual*
  - PP as fraction of total pay tends to *decline over second half of life cycle* (precisely when it should be *increasing*)
  - in fact, profile of PP relative to total pay is *hump-shaped* with experience in a variety of public and firm-level data sets

- Given this failure of existing models to account for basic features of data
  - how can we rationalize the pattern of PP over the life cycle?
  - do we need to account for performance incentives *at all* if PP on average accounts for no more than 10% of wages?

This Paper: We Address These Questions

- By first providing new evidence on importance of PP for life-cycle *w at odds* w/ predictions of existing models

This Paper: We Address These Questions

- By first providing new evidence on importance of PP for life-cycle *w at odds* w/ predictions of existing models
- Then reexamine role of incentives for wages by proposing new model
  - nests performance incentives, uncertainty and learning about ability, and human capital (HK) acquisition
  - micro-founds notion of jobs as requiring multiple tasks that are contractable to different degrees
  - allows us to decompose PP into the relative contribution of primitive forces we integrate

This Paper: We Address These Questions

- By first providing new evidence on importance of PP for life-cycle w *at odds* w/ predictions of existing models
- Then reexamine role of incentives for wages by proposing new model
  - nests performance incentives, uncertainty and learning about ability, and human capital (HK) acquisition
  - micro-founds notion of jobs as requiring multiple tasks that are contractable to different degrees
  - allows us to decompose PP into the relative contribution of primitive forces we integrate
- Show model
  - identified from panel data on wages and their fixed/variable components based on this characterization
  - reproduces dynamics of w, dispersion, their fixed-variable pay comp'n as well as workers' task assignments profiles

This Paper: We Address These Questions

- By first providing new evidence on importance of PP for life-cycle w *at odds* w/ predictions of existing models
- Then reexamine role of incentives for wages by proposing new model
  - nests performance incentives, uncertainty and learning about ability, and human capital (HK) acquisition
  - micro-founds notion of jobs as requiring multiple tasks that are contractable to different degrees
  - allows us to decompose PP into the relative contribution of primitive forces we integrate
- Show model
  - identified from panel data on wages and their fixed/variable components based on this characterization
  - reproduces dynamics of w, dispersion, their fixed-variable pay comp'n as well as workers' task assignments profiles
- Demonstrates model resolves puzzle on level and variability of PP
  - PP is *low* for insurance reasons to mitigate workers' *correlated life-cycle wage risk* due to uncertainty about their ability
  - eventually $\downarrow$ when effort to produce output augments HK (when HK less valuable, impl./expl. incentives optimally lower)
  - find PP central to dynamics of w: contributes $\approx 30\%$ of wage growth and variability over first 10 to 20 yrs of experience

Data

Data

- We use public worker panel data (PSID, NLSY79 and NLSY97) and confidential firm personnel records
  - from three influential studies in the literature (Baker-Gibbs-Holmström 1994a,b, Gibbs-Hendricks 2004)
  - to provide evidence on experience profile of wages and fixed/variable components

- In all samples: wages (labor earnings) are given by sum of fixed f_t and variable v_t pay: $w_t = f_t + v_t$

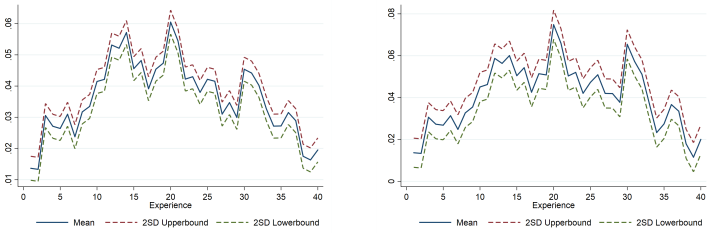
- Since in our model variable pay v_t proportional to performance via piece rate b_t: $v_t = b_t y_t$
  - can measure sensitivity of pay to performance b_t as $\mathbb{E}(v_t)/\mathbb{E}(w_t)$
  - under the assumption of free entry of firms in labor market since $\mathbb{E}(w_t) = \mathbb{E}(y_t)$

- Based on these data spanning across different years, workers and firms
  - we document ratio of PP to total eventually *declines* with experience
  - contrary to the prediction of CC models with explicit incentives

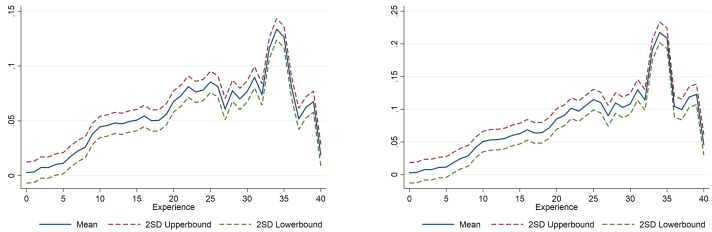
Next: for today will focus on our two firm-level data sets

In Both Data Experience Profile of Sensitivity Hump-Shaped

Managers and College-Educated Managers in BGH Sample



White-Collar Workers and College-Educated White-Collar Workers in GH Sample



Next: lay out model we propose to account for these patterns

Model

Competitive Labor Market

- Over finite horizon populated by homogeneous risk-neutral firms (can relax) and heterogeneous workers
  - have CARA preferences with parameter r, nonseparable over time $-\exp\{-r[\sum_{t=0}^{T} \delta^t (w_t - e_{1t}^2/2 - e_{2t}^2/2)]\}$
  - so workers indifferent among all deterministic wage streams with constant PV (as in Gibbons and Murphy, 1992)
    - note: cost of effort quadratic (second derivative set to 1 for ease of exposition and reasons of identification)

- Workers each period exert two types of effort
  - on *simple tasks* e_{1t} easy to monitor (*contractable*): entail creating/selling products or direct contacts with clients
  - on *complex tasks* e_{2t} difficult to monitor (*non-contractable*): entail managing large groups or strategic planning

  $\rightarrow$ we think of firms' *jobs* as primitive bundles of simple and complex tasks (continuum of them)

- Firms compete by offering one-period wage contracts linear in worker's output y_t
  - recall wage in any t, $w_t(y_t) = f_t + b_t y_t$, sum of fixed f_t and variable $v_t = b_t y_t$ pay
  - assuming long-term contracts infeasible equivalent to feasible but renegotiation-proof (reneg'ted if Pareto inefficient)

Output Technology, Ability and Information

- Worker (log) output at any t depends (log linearly) on worker ability, HK and effort and is subject to shocks
  - $y_t = \theta_t + \xi_k k_t + \xi_1 e_{1t} + \xi_2 e_{2t} + \varepsilon_t$
  - whereas ability θ_t is unobserved to all, human capital k_t and effort e_{2t} on *complex tasks* are observed *only* to a worker
  - e_{1t} on *simple tasks* is observable to all

Output Technology, Ability and Information

- Worker (log) output at any t depends (log linearly) on worker ability, HK and effort and is subject to shocks
  - $y_t = \theta_t + \xi_k k_t + \xi_1 e_{1t} + \xi_2 e_{2t} + \varepsilon_t$
  - whereas ability θ_t is unobserved to all, human capital k_t and effort e_{2t} on *complex tasks* are observed *only* to a worker
  - e_{1t} on *simple tasks* is observable to all

- $\theta_{t+1} = \theta_t + \zeta_t$: ability θ_t evolves over time according to random walk process
  - θ_0: normally distributed w/ mean zero and variance σ_θ^2
  - ζ_t: mean-zero normal shock w/ variance σ_ζ^2 for unanticipated variations in θ_t (can also be interpreted as shock to HK)
  - differently from CC models: allow for renewal of uncertainty about θ_t through ζ_t

Output Technology, Ability and Information

- Worker (log) output at any t depends (log linearly) on worker ability, HK and effort and is subject to shocks
  - $y_t = \theta_t + \xi_k k_t + \xi_1 e_{1t} + \xi_2 e_{2t} + \varepsilon_t$
  - whereas ability θ_t is unobserved to all, human capital k_t and effort e_{2t} on *complex tasks* are observed *only* to a worker
  - e_{1t} on *simple tasks* is observable to all

- $\theta_{t+1} = \theta_t + \zeta_t$: ability θ_t evolves over time according to random walk process
  - θ_0: normally distributed w/ mean zero and variance σ_θ^2
  - ζ_t: mean-zero normal shock w/ variance σ_ζ^2 for unanticipated variations in θ_t (can also be interpreted as shock to HK)
  - differently from CC models: allow for renewal of uncertainty about θ_t through ζ_t

- At end of each period t: beliefs about θ_t are updated based on realized y_t according to Bayes' rule
  - given worker's conjectured $\hat{k}_t$ and $(\hat{e}_{1t}\ \hat{e}_{2t})$: $z_t = y_t - \xi_k \hat{k}_t - \xi_1 \hat{e}_{1t} - \xi_2 \hat{e}_{2t}$ is signal about θ_t extracted from y_t
  - so workers have an incentive to exert effort to affect signal z_t and so market beliefs about θ_t (CC incentive for effort)

Human Capital Technology

- $k_{t+1} = \lambda k_t + \gamma_1(e_{1t} - \overline{e}_t) + \gamma_2(e_{2t} - \overline{e}_t)$: HK accumulates with effort in each task at rates (γ_1, γ_2)
  - depreciates at rate $1 - \lambda$ (can allow for semiparametric law $k_{t+1} = \lambda k_t + F(e_{1t}, e_{2t})$ with performance information)

Human Capital Technology

- $k_{t+1} = \lambda k_t + \gamma_1(e_{1t} - \overline{e}_t) + \gamma_2(e_{2t} - \overline{e}_t)$: HK accumulates with effort in each task at rates (γ_1, γ_2)
  - depreciates at rate $1 - \lambda$ (can allow for semiparametric law $k_{t+1} = \lambda k_t + F(e_{1t}, e_{2t})$ with performance information)

- HK can be acquired through learning-by-doing (LBD) and learning-or-doing (Ben-Porath or LOD)
  - when $\gamma_j > 0$: effort to produce output in task j (e_{jt}) also produces HK so the two efforts are *complements* (LBD)
  - when $\gamma_j < 0$: effort to produce output in task j (e_{jt}) and to produce HK $(\overline{e}_t - e_t)$ are *substitutes* (rival: LOD)

Human Capital Technology

- $k_{t+1} = \lambda k_t + \gamma_1(e_{1t} - \overline{e}_t) + \gamma_2(e_{2t} - \overline{e}_t)$: HK accumulates with effort in each task at rates (γ_1, γ_2)
  - depreciates at rate $1 - \lambda$ (can allow for semiparametric law $k_{t+1} = \lambda k_t + F(e_{1t}, e_{2t})$ with performance information)

- HK can be acquired through learning-by-doing (LBD) and learning-or-doing (Ben-Porath or LOD)
  - when $\gamma_j > 0$: effort to produce output in task j (e_{jt}) also produces HK so the two efforts are *complements* (LBD)
  - when $\gamma_j < 0$: effort to produce output in task j (e_{jt}) and to produce HK ($\overline{e}_t - e_t$) are *substitutes* (rival: LOD)

- Note: our specification of worker efficient labor supply as $\theta_t + g(t) + \xi_1 e_{1t} + \xi_2 e_{2t} + \varepsilon_t$ with $g(t) \equiv k_t$
  - is extension of one common in literature (e.g. Bagger et al. 2014: $\theta + g(t) + \varepsilon_t$)
  - main differences: we allow worker heterogeneity parameter θ_t to be unknown and time varying
  - efforts in t to affect amount of efficient labor provided and HK to evolve *endogenously* as function of past efforts

Model Nests and Extends Several Known Models

- In literature on MH, learning about ability and HK accumulation

- In particular, when HK acquisition and shocks to ability muted ($\gamma_t = \sigma^2_\zeta = 0$) and there is no contractible effort
  - it reduces to the CC model w/ explicit contracts of Gibbons and Murphy (1992)
  - if in addition explicit contracts are infeasible: it is standard CC model of Holmström (1982, 1999)
  - if further effort is fixed: specializes to typical learning model w/ general ability as Farber and Gibbons (1996)

- When instead ability is known and all effort if contractible
  - model reduces to one of *labor supply and HK acquisition* through LBD or LOD
  - if in addition effort fixed: model is of HK acquisition with experience

Next: turn to characterize sensitivity of pay to performance

Sensitivity of Pay to Performance Implied by Model

- Given by $b_t^* = b_t^0 \left[1 - \underbrace{R^*_{CC,t}}_{(1)} - \underbrace{rH^*_t}_{(2)} + \underbrace{\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}}_{(3)} \right]$ where
  - scaling factor $b_t^0 = 1/[1 + r(\sigma_t^2 + \sigma_\varepsilon^2)]$ is standard piece rate from static (linear normal) MH models
  - so optimal b_t^* differs from static one due to last three terms
    - first two are negative so depress piece rates relative to static level
    - last has ambiguous effect: consists of positive and negative term

- Note that if workers were risk neutral, $b_t^* = 1$ so v_t would move 1-1 with y_t
  - since workers are risk averse, firms optimally provide incentives by smoothing variability of w over time
  - they do so balancing at the margin strength of explicit and implicit incentives

Sensitivity of Pay to Performance Implied by Model

- Given by $b_t^* = b_t^0 \left[1 - \underbrace{R^*_{CC,t}}_{(1)} - \underbrace{rH^*_t}_{(2)} + \underbrace{\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}}_{(3)} \right]$ where
  - scaling factor $b_t^0 = 1/[1 + r(\sigma_t^2 + \sigma_\varepsilon^2)]$ is standard piece rate from static (linear normal) MH models
  - so optimal b_t^* differs from static one due to last three terms
    - first two are negative so depress piece rates relative to static level
    - last has ambiguous effect: consists of positive and negative term

- Note that if workers were risk neutral, $b_t^* = 1$ so v_t would move 1-1 with y_t
  - since workers are risk averse, firms optimally provide incentives by smoothing variability of w over time
  - they do so balancing at the margin strength of explicit and implicit incentives

- $R^*_{CC,t} = \sum_{\tau=1}^{T-t} \delta^\tau (1 - b^*_{t+\tau}) \partial \mathbb{E}_t(\theta_{t+\tau}) / \partial e_{2t}$: CC provide implicit effort incentives even in absence of PP
  - so by partially substituting for explicit incentives lead to lower piece rates

Sensitivity of Pay to Performance Implied by Model

Given by $b_t^* = b_t^0 \left[1 - \underbrace{R^*_{CC,t}}_{(1)} - \underbrace{rH^*_t}_{(2)} + \underbrace{\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}}_{(3)} \right]$ where

- $rH^*_t = r\sigma_t^2 \sum_{\tau=1}^{T-t} \delta^\tau$: captures insurance against income risk due to uncertainty via $\downarrow$ piece rates
  - $\propto$ variance of beliefs which leads to variability in w_t workers dislike (firms partially shield workers $\downarrow corr(w_t, w_{t+\tau})$)
  - term depends on r: optimal insurance higher, the more risk averse workers are

Sensitivity of Pay to Performance Implied by Model

Given by $b_t^* = b_t^0 \left[1 - \underbrace{R^*_{CC,t}}_{(1)} - \underbrace{rH_t^*}_{(2)} + \underbrace{\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}}_{(3)} \right]$ where

- $rH_t^* = r\sigma_t^2 \sum_{\tau=1}^{T-t} \delta^\tau$: captures insurance against income risk due to uncertainty via $\downarrow$ piece rates
  - $\propto$ variance of beliefs which leads to variability in w_t workers dislike (firms partially shield workers $\downarrow corr(w_t, w_{t+\tau})$)
  - term depends on r: optimal insurance higher, the more risk averse workers are

- $\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}$: novel adjustment term for HK consists of two further components

Sensitivity of Pay to Performance Implied by Model

Given by $b_t^* = b_t^0 \left[1 - \underbrace{R^*_{CC,t}}_{(1)} - \underbrace{rH^*_t}_{(2)} + \underbrace{\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}}_{(3)} \right]$ where

- $rH^*_t = r\sigma_t^2 \sum_{\tau=1}^{T-t} \delta^\tau$: captures insurance against income risk due to uncertainty via $\downarrow$ piece rates
  - $\propto$ variance of beliefs which leads to variability in w_t workers dislike (firms partially shield workers $\downarrow corr(w_t, w_{t+\tau})$)
  - term depends on r: optimal insurance higher, the more risk averse workers are

- $\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}$: novel adjustment term for HK consists of two further components
  - first: accounts for impact of effort on future output (social marginal return)

Sensitivity of Pay to Performance Implied by Model

Given by $b_t^* = b_t^0 \left[1 - \underbrace{R^*_{CC,t}}_{(1)} - \underbrace{rH^*_t}_{(2)} + \underbrace{\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}}_{(3)} \right]$ where

- $rH^*_t = r\sigma_t^2 \sum_{\tau=1}^{T-t} \delta^\tau$: captures insurance against income risk due to uncertainty via ↓ piece rates
  - $\propto$ variance of beliefs which leads to variability in w_t workers dislike (firms partially shield workers ↓ $corr(w_t, w_{t+\tau})$)
  - term depends on r: optimal insurance higher, the more risk averse workers are

- $\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}$: novel adjustment term for HK consists of two further components
  - first: accounts for impact of effort on future output (social marginal return)
  - second: accounts for HK implicitly incentivizing e_{2t} (private marginal return) so partially substitutes for explicit incent's
    - b_t^* provides incentives just for the portion of HK return that workers do not internalize

Sensitivity of Pay to Performance Implied by Model

Given by $b_t^* = b_t^0 \left[1 - \underbrace{R^*_{CC,t}}_{(1)} - \underbrace{rH^*_t}_{(2)} + \underbrace{\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}}_{(3)} \right]$ where

- $rH^*_t = r\sigma_t^2 \sum_{\tau=1}^{T-t} \delta^\tau$: captures insurance against income risk due to uncertainty via ↓ piece rates
  - $\propto$ variance of beliefs which leads to variability in w_t workers dislike (firms partially shield workers ↓ $corr(w_t, w_{t+\tau})$)
  - term depends on r: optimal insurance higher, the more risk averse workers are

- $\gamma_t \sum_{\tau=1}^{T-1} \delta^\tau \lambda^{\tau-1} - R^*_{LBD,t}$: novel adjustment term for HK consists of two further components
  - first: accounts for impact of effort on future output (social marginal return)
  - second: accounts for HK implicitly incentivizing e_{2t} (private marginal return) so partially substitutes for explicit incent's
    - b_t^* provides incentives just for the portion of HK return that workers do not internalize

Next: experience profile of $\{b_t^*\}$ naturally depends on relative strength of all these forces

Alternative Profiles of Sensitivity of Pay to Performance Possible

- For instance, with ability uncertainty but w/o HK (or with LOD HK): $b_t^* \uparrow$ with t as in Gibbons-Murphy (1992)
  - intuition: explicit incentives substitute for implicit ones as uncertainty decreases

Alternative Profiles of Sensitivity of Pay to Performance Possible

- For instance, with ability uncertainty but w/o HK (or with LOD HK): $b_t^* \uparrow$ with t as in Gibbons-Murphy (1992)
  - intuition: explicit incentives substitute for implicit ones as uncertainty decreases

- W/o ability uncertainty but with LBD HK, piece rates exhibit opposite profile: $b_t^* \downarrow$ with t
  - intuition: returns to HK investments decline over time as residual lifetime shortens
  - so optimal to provide strongest incentives early on (i.e. explicit and implicit incentives complements)

Alternative Profiles of Sensitivity of Pay to Performance Possible

- For instance, with ability uncertainty but w/o HK (or with LOD HK): $b_t^* \uparrow$ with t as in Gibbons-Murphy (1992)
  - intuition: explicit incentives substitute for implicit ones as uncertainty decreases

- W/o ability uncertainty but with LBD HK, piece rates exhibit opposite profile: $b_t^* \downarrow$ with t
  - intuition: returns to HK investments decline over time as residual lifetime shortens
  - so optimal to provide strongest incentives early on (i.e. explicit and implicit incentives complements)

- Combined model inherits both forces so leads to
  - hump-shaped profile of b_t^*: if LBD HK motives *weaker* than uncertainty at low levels of experience
    - e.g. when σ_θ^2 is large, σ_ζ^2 is small and T is large enough
  - u-shaped profile of b_t^*: if LBD HK motives *stronger* than uncertainty at low levels of experience
    - when LBD HK accumulation rates not too small and the signal-to-noise ratio is large enough

Alternative Profiles of Sensitivity of Pay to Performance Possible

- For instance, with ability uncertainty but w/o HK (or with LOD HK): $b_t^* \uparrow$ with t as in Gibbons-Murphy (1992)
  - intuition: explicit incentives substitute for implicit ones as uncertainty decreases

- W/o ability uncertainty but with LBD HK, piece rates exhibit opposite profile: $b_t^* \downarrow$ with t
  - intuition: returns to HK investments decline over time as residual lifetime shortens
  - so optimal to provide strongest incentives early on (i.e. explicit and implicit incentives complements)

- Combined model inherits both forces so leads to
  - hump-shaped profile of b_t^*: if LBD HK motives *weaker* than uncertainty at low levels of experience
    - e.g. when σ_θ^2 is large, σ_ζ^2 is small and T is large enough
  - u-shaped profile of b_t^*: if LBD HK motives *stronger* than uncertainty at low levels of experience
    - when LBD HK accumulation rates not too small and the signal-to-noise ratio is large enough

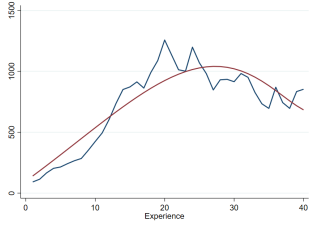
- Based on these arguments, prove model identified
  - key intuition: equilibrium b_t^* provides known mapping btw PP and worker preference/HK parameters (as we saw)

Estimation Results

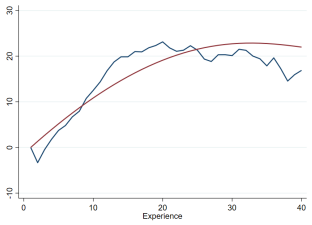
Estimates of Model on Firm-Level Data From BGH

- With fixed δ: 9 remaining parameters $(\sigma_\theta^2, \sigma_\zeta^2, \sigma_\varepsilon^2, \gamma_1, \gamma_2, \lambda, r)$ estimated by MD targeting 120 moments
- Corresponding to profile of w variance, cumulative w growth and b_t^* over first 40 years of experience

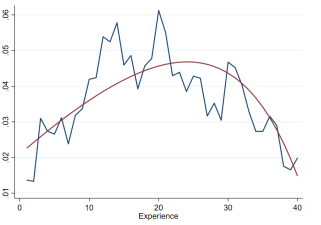
(a) Wage Variance



(b) Wage Growth



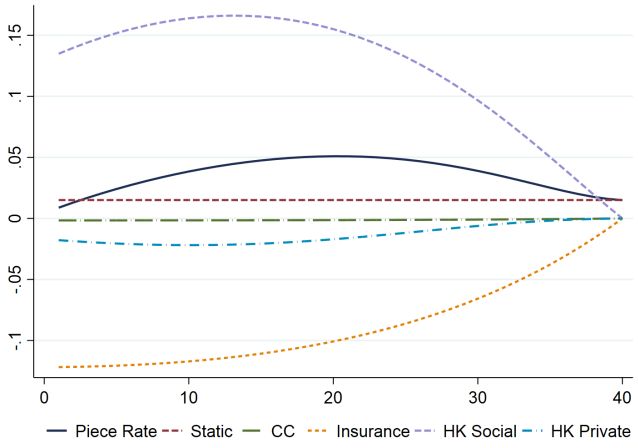
(c) Piece Rates



- As apparent from figure: model closely matches all these dimensions of data (very precisely)
- Based on estimates: can decompose estimated b_t^* into the five components isolated earlier

Decompose Estimated Piece Rate at Each Experience

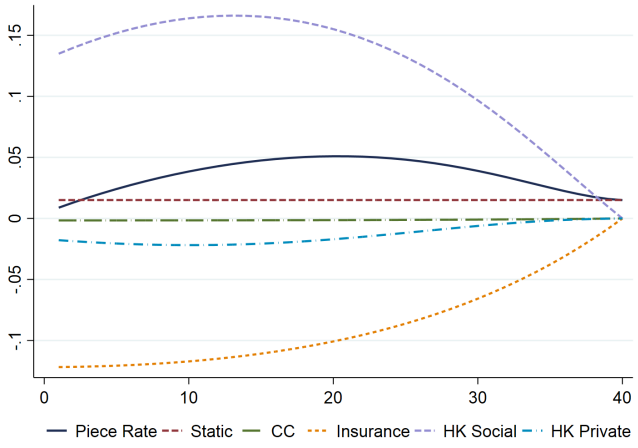
Into components due to: static piece rate, CC, insurance, HK social and HK private return



Find key components: HK social and insurance vs. uncertainty about θ_t (figure: remaining negligible)

Decompose Estimated Piece Rate at Each Experience

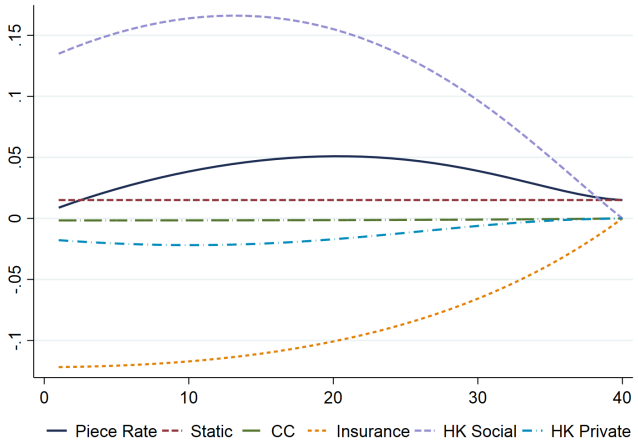
Into components due to: static piece rate, CC, insurance, HK social and HK private return



HK social, which is large and positive, important to account for hump-shape of piece rates

Decompose Estimated Piece Rate at Each Experience

Into components due to: static piece rate, CC, insurance, HK social and HK private return



Insurance against uncertainty about θ_t, which is fairly large and negative term, explains low level

Can Interpret Decomposition in Light of Estimates

- A robust findings of different parameterizations of our model (e.g. imposing fast or slow learning)
  - output shocks must be large relative to uncertainty about ability for b_t^* and variance of w to be $\downarrow$ later in life
  - as apparent from $\text{Var}[w_{it}] = \sigma_\theta^2 + t\sigma_\zeta^2 - \sigma_t^2 + (b_t^*)^2(\sigma_t^2 + \sigma_\varepsilon^2)$ (hump shape in b_t^* translates into one in $\text{Var}[w_{it}]$)

Can Interpret Decomposition in Light of Estimates

- A robust findings of different parameterizations of our model (e.g. imposing fast or slow learning)
  - output shocks must be large relative to uncertainty about ability for b_t^* and variance of w to be $\downarrow$ later in life
  - as apparent from $\text{Var}[w_{it}] = \sigma_\theta^2 + t\sigma_\zeta^2 - \sigma_t^2 + (b_t^*)^2(\sigma_t^2 + \sigma_\varepsilon^2)$ (hump shape in b_t^* translates into one in $\text{Var}[w_{it}]$)
    - this explains why effort has small impact on beliefs thus CC incentives small (limited impact of effort on signals)
    - why static piece rate $b_t^0 = 1/[1 + r(\sigma_t^2 + \sigma_\varepsilon^2)]$ is low so are piece rates
    - why private marginal return to HK small since $\propto$ to CC incentives and piece rates (both small)

- Although estimate absolute risk aversion moderate and σ_θ^2 not too large: insurance component large

Can Interpret Decomposition in Light of Estimates

- A robust findings of different parameterizations of our model (e.g. imposing fast or slow learning)
  - output shocks must be large relative to uncertainty about ability for b_t^* and variance of w to be $\downarrow$ later in life
  - as apparent from $\text{Var}[w_{it}] = \sigma_\theta^2 + t\sigma_\zeta^2 - \sigma_t^2 + (b_t^*)^2(\sigma_t^2 + \sigma_\varepsilon^2)$ (hump shape in b_t^* translates into one in $\text{Var}[w_{it}]$)
    - this explains why effort has small impact on beliefs thus CC incentives small (limited impact of effort on signals)
    - why static piece rate $b_t^0 = 1/[1 + r(\sigma_t^2 + \sigma_\varepsilon^2)]$ is low so are piece rates
    - why private marginal return to HK small since $\propto$ to CC incentives and piece rates (both small)

- Although estimate absolute risk aversion moderate and σ_θ^2 not too large: insurance component large. Why?

Can Interpret Decomposition in Light of Estimates

- A robust findings of different parameterizations of our model (e.g. imposing fast or slow learning)
  - output shocks must be large relative to uncertainty about ability for b_t^* and variance of w to be $\downarrow$ later in life
  - as apparent from $\text{Var}[w_{it}] = \sigma_\theta^2 + t\sigma_\zeta^2 - \sigma_t^2 + (b_t^*)^2(\sigma_t^2 + \sigma_\varepsilon^2)$ (hump shape in b_t^* translates into one in $\text{Var}[w_{it}]$)
    - this explains why effort has small impact on beliefs thus CC incentives small (limited impact of effort on signals)
    - why static piece rate $b_t^0 = 1/[1 + r(\sigma_t^2 + \sigma_\varepsilon^2)]$ is low so are piece rates
    - why private marginal return to HK small since $\propto$ to CC incentives and piece rates (both small)

- Although estimate absolute risk aversion moderate and σ_θ^2 not too large: insurance component large. Why?

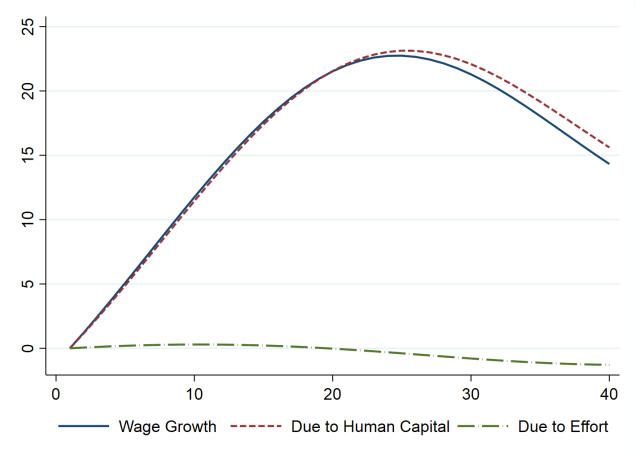
- Intuition simple from asset pricing perspective: v_t high when y_t high so news about θ_t and future w positive
  - so contracts with v_t pay out more claims to y_t in good times when MU_C low but less in bad times when MU_C high
  - thus exhibit *opposite covariance* structure of returns than risk-averse investors desire
  - low $b_t^* \downarrow$ this risk by $\downarrow$ correlation btw current and future pay (insurance against life-cycle risk from uncertainty)

Next: discuss role of performance incentives for wage growth ($w_t - w_1$ in thousands \$) and dispersion

What Accounts for Lifecycle Wage Growth?

As growth is sum growth in HK and effort: can decompose in *direct* contribution of each (figure)

What Accounts for Lifecycle Wage Growth?

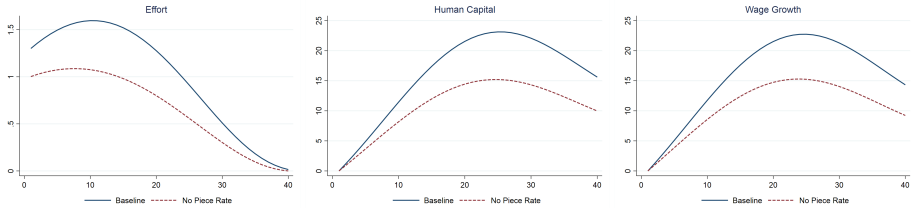


By performing this decomposition: find HK accounts for almost all growth (red vs. blue)

Decomposition Masks Impact Effort on Human Capital

But effort has important *indirect* effect on w growth as active margin of investment in HK (find it of the LBD type)

One way to see how e_{2t} matters: assume firms restricted to offer contracts w/o variable pay

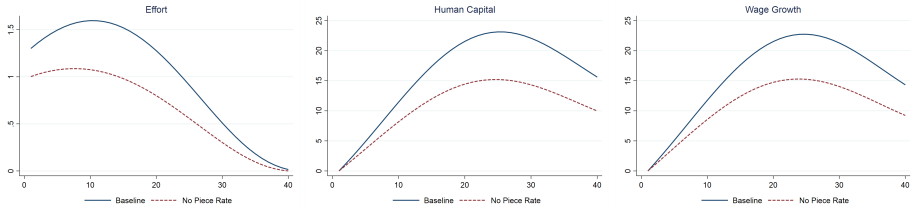


W/o p-incentives ($b_t^* = 0$): e_{2t} and k_t much lower would lead to 30% ↓ growth (see red vs. blue)

Decomposition Masks Impact Effort on Human Capital

But effort has important *indirect* effect on w growth as active margin of investment in HK (find it of the LBD type)

One way to see how e_{2t} matters: assume firms restricted to offer contracts w/o variable pay

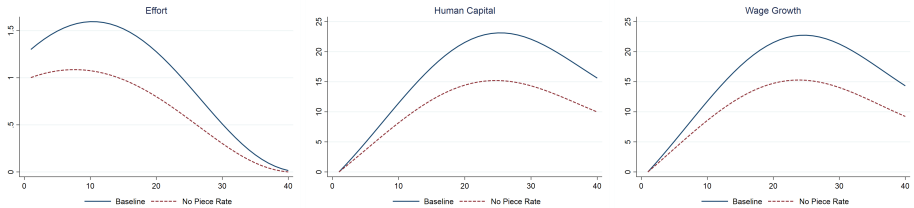


So performance incentives *matter* for wage growth

Decomposition Masks Impact Effort on Human Capital

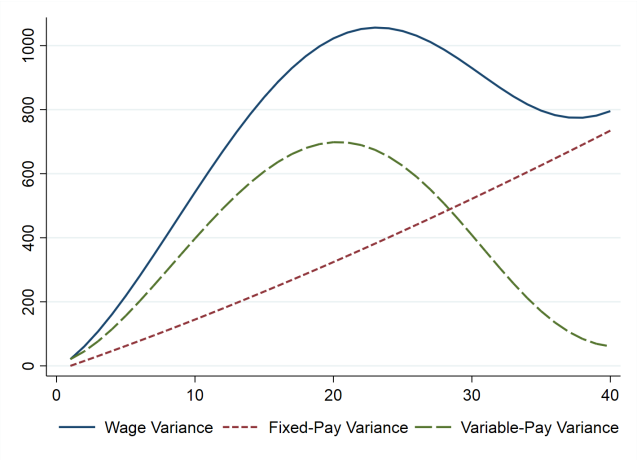
But effort has important *indirect* effect on w growth as active margin of investment in HK (find it of the LBD type)

One way to see how e_{2t} matters: assume firms restricted to offer contracts w/o variable pay



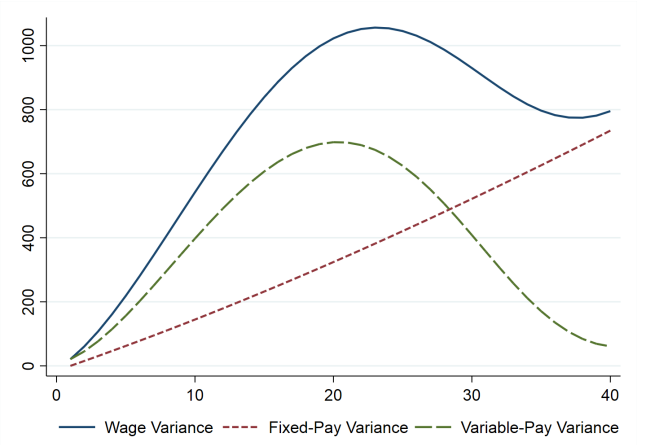
So performance incentives *matter* for wage growth. Do they matter for wage dispersion?

Although Small Variable Pay Also Key to Variance of Wages



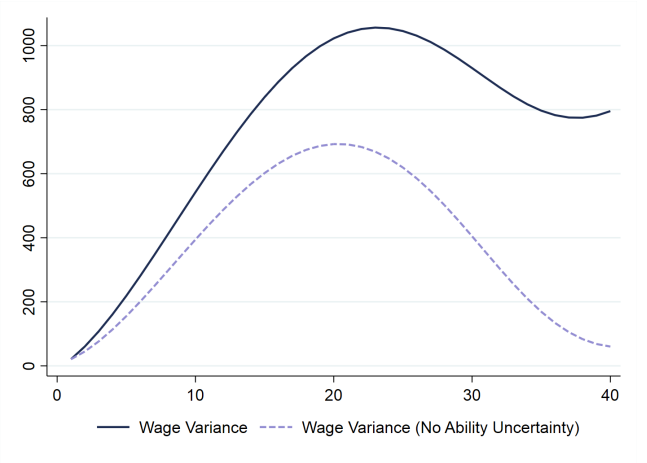
Over lifecycle: as apparent from decomposing it into contribution of fixed and variable components of pay (figure)

Although Small Variable Pay Also Key to Variance of Wages



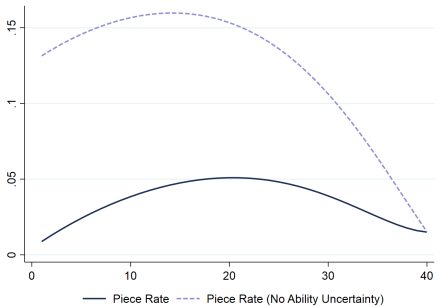
Indeed variance v_t alone (green) accounts for no less than 35% of variance w_t (blue) over first 20 years

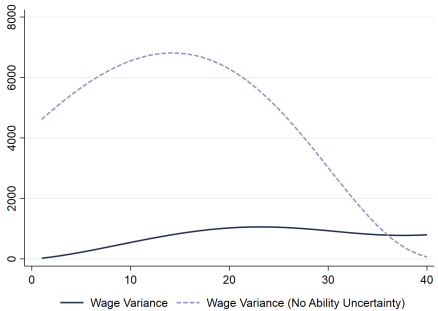
How Important Is Ability Uncertainty for Wage Dispersion?



At the estimated equilibrium $\{b_t^*\}$: large fraction of w variance accounted by it $(\sigma_\theta^2 = \sigma_\zeta^2 = 0)$

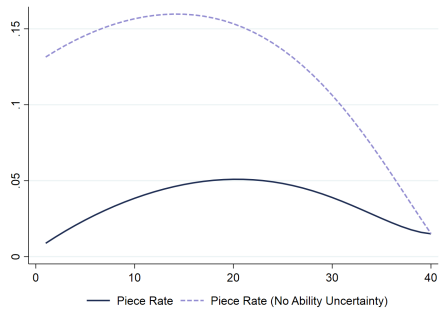
But Lowering Ability Uncertainty Would Not Lower Dispersion

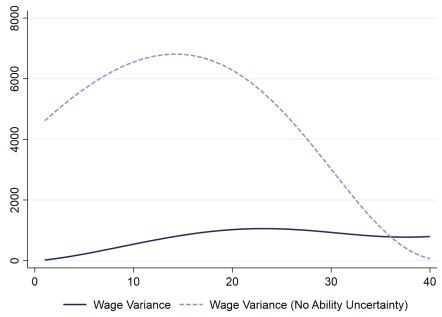




Why? Workers demand less insurance so firms offer higher b_t^*: amplify risk leading to *much* ↑ variability

But Lowering Ability Uncertainty Would Not Lower Dispersion





W/o uncertainty about θ_t: piece rates up to 3 times higher and variance up to 4 times larger

What Do We Learn From This Experiment?

- Much evidence points to heterogeneity in ability as important source of persistent w inequality among workers
- But wage structure is itself determined by this heterogeneity
- In particular $\downarrow$ it can actually lead to $\uparrow$ w dispersion by inducing firms to offer wages more sensitive to y_t
- So mitigating skill differences at entry (e.g. better schooling) can end up amplifying wage inequality
- Key idea: wage structure is important endogenous transmission mechanism of shocks to wages

What About Task Complexity over Life Cycle?

- So far focused only on non-contractable effort

What About Task Complexity over Life Cycle?

- So far focused only on non-contractable effort: once we allow for contractable effort
  - the model better fits average wage growth with the additional parameter γ_1 to match it
  - but also find γ_2 is only slightly lower
  - so incorporating standard dimension of labor supply does not decrease much importance of performance incentives for w

What About Task Complexity over Life Cycle?

- So far focused only on non-contractable effort: once we allow for contractable effort
  - the model better fits average wage growth with the additional parameter γ_1 to match it
  - but also find γ_2 is only slightly lower
  - so incorporating standard dimension of labor supply does not decrease much importance of performance incentives for w

- We also find that although effort paths implied by our model are eventually *declining* over time
  - they are characterized by an approximately *increasing* degree of task complexity $(1 + e^*_{2t})/(1 + e^*_{1t})$
  - this is consistent w/ importance of general manag't (complex tasks) vs. product/client act's (simple tasks) in BGH data
  - according to BGH' definition of complexity of jobs and their description of jobs' task content

- These results validate the incentives, human capital and uncertainty mechanisms at the heart of our model

Conclusion

- We have proposed new model of learning, HK and performance incentives to account for
  - overall level of wages, their dispersion, their composition in terms of fixed and variable pay and dynamics
  - which rationalizes puzzle that ratio of variable to total pay *declines* over second half of life cycle

- We have characterized optimal wage contract and based on this characterization
  - isolated the distinct determinants of the level and experience profile of PP relative to total pay
  - proved model identified just from panel data on w and their fixed or variable components

- Our estimation results show
  - insurance against life-cycle wage risk arising from uncertainty about ability is main reason for low PP
  - yet performance incentives key to dynamics of overall wages both directly and through impact on workers' HK process

- Hope first step toward richer models of incentives to interpret sources of dispersion in wages and its persistence