# On the Identification of Models of Uncertainty, Learning, and Human Capital Acquisition with Sorting

Aureo De Paula[1] Cristina Gualdani[2] Elena Pastorino[3] Sergio Salgado[4]

[1]UCL [2]Queen Mary [3]Stanford and Hoover [4]Wharton

DSE Conference 2025

# Motivation

- *Canonical labor market models* interpret worker mobility, wage growth and wage dispersion as driven by
  - ▷ *Dynamic matching* process between firms and workers
  - ▷ *Human capital* (HK) *acquisition* through labor market experience (Becker, 1962; Mincer, 1974)
  - ▷ Gradual *learning* about a worker's true *productivity* (Jovanovic, 1979)

- Open question whether this workhorse class of models is *identified*: difficult to establish
  - ▷ since these forces lead to dynamic equilibrium generalized Roy models with selection on unobservables

- Yet answer is key because these models represent classic framework to study *sorting* in labor market

# Motivation

- *Canonical labor market models* interpret worker mobility, wage growth and wage dispersion as driven by
  - ▷ *Dynamic matching* process between firms and workers
  - ▷ *Human capital* (HK) *acquisition* through labor market experience (Becker, 1962; Mincer, 1974)
  - ▷ Gradual *learning* about a worker's true *productivity* (Jovanovic, 1979)

- Open question whether this workhorse class of models is *identified*: difficult to establish
  - ▷ since these forces lead to dynamic equilibrium generalized Roy models with selection on unobservables

- Yet answer is key because these models represent classic framework to study *sorting* in labor market

# Motivation

- *Canonical labor market models* interpret worker mobility, wage growth and wage dispersion as driven by
  - ▷ *Dynamic matching* process between firms and workers
  - ▷ *Human capital* (HK) *acquisition* through labor market experience (Becker, 1962; Mincer, 1974)
  - ▷ Gradual *learning* about a worker's true *productivity* (Jovanovic, 1979)

- Open question whether this workhorse class of models is *identified*: difficult to establish
  - ▷ since these forces lead to dynamic equilibrium generalized Roy models with selection on unobservables

- Yet answer is key because these models represent classic framework to study *sorting* in labor market

# Motivation

- Indeed, wage inequality is typically measured through empirical models predicated on sorting

- But impact of sorting on wage inequality is usually estimated to be low (e.g. AKM, Card & al. (2013))
  - ▷ Raises puzzle: given the large degree of wage inequality (especially in U.S.)
  - ▷ Why is sorting estimated to be *unimportant* if canonical models of inequality are based on it?

- Our answer: dynamic matching models with HK acquisition and learning lead to two *confounding* forces

1. Naturally give rise to *countervailing effects (compensating differential)* in the wage equation
   - ▷ Compensate worker for lost opportunity of HK acquisition and learning at *other* competing firms
   - ▷ Attenuate impact of firm/worker types on wages contaminating traditional measures of sorting

2. Feature endogenous matching *"frictions"*: as *HK acquisition*/*learning* about ability take place over time
   - ▷ Workers may temporarily choose less productive firms offering valuable training/learning opportunities
   - ▷ So a *highly productive* worker can be paid a relatively *low wage*

# Motivation

- Indeed, wage inequality is typically measured through empirical models predicated on sorting
- But impact of sorting on wage inequality is usually estimated to be low (e.g. AKM, Card & al. (2013))
  - ▷ Raises puzzle: given the large degree of wage inequality (especially in U.S.)
  - ▷ Why is sorting estimated to be *unimportant* if canonical models of inequality are based on it?
- Our answer: dynamic matching models with HK acquisition and learning lead to two *confounding* forces

1. Naturally give rise to *countervailing effects (compensating differential)* in the wage equation
   - ▷ Compensate worker for lost opportunity of HK acquisition and learning at *other* competing firms
   - ▷ Attenuate impact of firm/worker types on wages contaminating traditional measures of sorting
2. Feature endogenous matching *"frictions"*: as *HK acquisition/learning* about ability take place over time
   - ▷ Workers may temporarily choose less productive firms offering valuable training/learning opportunities
   - ▷ So a *highly productive* worker can be paid a relatively *low wage*

# Motivation

- Indeed, wage inequality is typically measured through empirical models predicated on sorting
- But impact of sorting on wage inequality is usually estimated to be low (e.g. AKM, Card & al. (2013))
  - ▷ Raises puzzle: given the large degree of wage inequality (especially in U.S.)
  - ▷ Why is sorting estimated to be *unimportant* if canonical models of inequality are based on it?
- Our answer: dynamic matching models with HK acquisition and learning lead to two *confounding* forces

1. Naturally give rise to *countervailing effects (compensating differential)* in the wage equation
   - ▷ Compensate worker for lost opportunity of HK acquisition and learning at *other* competing firms
   - ▷ Attenuate impact of firm/worker types on wages contaminating traditional measures of sorting
2. Feature endogenous matching *"frictions"*: as *HK acquisition/learning* about ability take place over time
   - ▷ Workers may temporarily choose less productive firms offering valuable training/learning opportunities
   - ▷ So a *highly productive* worker can be paid a relatively *low wage*

# Motivation

- Indeed, wage inequality is typically measured through empirical models predicated on sorting
- But impact of sorting on wage inequality is usually estimated to be low (e.g. AKM, Card & al. (2013))
  - ▷ Raises puzzle: given the large degree of wage inequality (especially in U.S.)
  - ▷ Why is sorting estimated to be *unimportant* if canonical models of inequality are based on it?
- Our answer: dynamic matching models with HK acquisition and learning lead to two *confounding* forces

1. Naturally give rise to *countervailing effects (compensating differential)* in the wage equation
   - ▷ Compensate worker for lost opportunity of HK acquisition and learning at *other* competing firms
   - ▷ Attenuate impact of firm/worker types on wages contaminating traditional measures of sorting
2. Feature endogenous matching *"frictions"*: as *HK acquisition*/*learning* about ability take place over time
   - ▷ Workers may temporarily choose less productive firms offering valuable training/learning opportunities
   - ▷ So a *highly productive* worker can be paid a relatively *low wage*

# Motivation

- Indeed, wage inequality is typically measured through empirical models predicated on sorting
- But impact of sorting on wage inequality is usually estimated to be low (e.g. AKM, Card & al. (2013))
  - ▷ Raises puzzle: given the large degree of wage inequality (especially in U.S.)
  - ▷ Why is sorting estimated to be *unimportant* if canonical models of inequality are based on it?
- Our answer: dynamic matching models with HK acquisition and learning lead to two *confounding* forces

1. Naturally give rise to *countervailing effects (compensating differential)* in the wage equation
   - ▷ Compensate worker for lost opportunity of HK acquisition and learning at *other* competing firms
   - ▷ Attenuate impact of firm/worker types on wages contaminating traditional measures of sorting
2. Feature endogenous matching *"frictions"*: as *HK acquisition/learning* about ability take place over time
   - ▷ Workers may temporarily choose less productive firms offering valuable training/learning opportunities
   - ▷ So a *highly productive* worker can be paid a relatively *low wage*

# Challenges to Identification

- Framework we consider features rich dynamic selection on unobservables
  - ▷ Wages hence employment depend on time-varying, serially corr. and endogenously evolving unobservables
- Also, unobservables affect wages in potentially nonmonotonic, nonseparable and nonmultiplicative manner
  - ▷ Compensating differential is difference in v-functions (end. dyn. payoffs) rather than per-period payoffs
  - ▷ Thus, interactive fixed effect approach is infeasible
- No exclusion restrictions arise from theory to solve selection (unlike in most Roy settings)

Bottom line: not immediate to adapt existing econometric approaches to establish identification

# Challenges to Identification

- Framework we consider features rich dynamic selection on unobservables
  - ▷ Wages hence employment depend on time-varying, serially corr. and endogenously evolving unobservables
- Also, unobservables affect wages in potentially nonmonotonic, nonseparable and nonmultiplicative manner
  - ▷ Compensating differential is difference in v-functions (end. dyn. payoffs) rather than per-period payoffs
  - ▷ Thus, interactive fixed effect approach is infeasible
- No exclusion restrictions arise from theory to solve selection (unlike in most Roy settings)

Bottom line: not immediate to adapt existing econometric approaches to establish identification

# Challenges to Identification

- Framework we consider features rich dynamic selection on unobservables
  - ▷ Wages hence employment depend on time-varying, serially corr. and endogenously evolving unobservables

- Also, unobservables affect wages in potentially nonmonotonic, nonseparable and nonmultiplicative manner
  - ▷ Compensating differential is difference in v-functions (end. dyn. payoffs) rather than per-period payoffs
  - ▷ Thus, interactive fixed effect approach is infeasible

- No exclusion restrictions arise from theory to solve selection (unlike in most Roy settings)

Bottom line: not immediate to adapt existing econometric approaches to establish identification

# Challenges to Identification

- Framework we consider features rich dynamic selection on unobservables
  - ▷ Wages hence employment depend on time-varying, serially corr. and endogenously evolving unobservables

- Also, unobservables affect wages in potentially nonmonotonic, nonseparable and nonmultiplicative manner
  - ▷ Compensating differential is difference in v-functions (end. dyn. payoffs) rather than per-period payoffs
  - ▷ Thus, interactive fixed effect approach is infeasible

- No exclusion restrictions arise from theory to solve selection (unlike in most Roy settings)

Bottom line: not immediate to adapt existing econometric approaches to establish identification

# What We Do: Two Sets of Results

1. Develop semiparametric identification arguments for matching models with HK acquisition and learning
   - ▷ Represent wage distribution as a mixture over unobservables
   - ▷ Establish identification of this wage mixture under nonparametric restrictions
   - ▷ Recover primitives by integrating mixture-based approach w/ quantile methods for standard Roy models
2. Use it to reevaluate impact of worker-firm sorting on U.S. earnings inequality using LEHD data
   - ▷ Findings: sorting matters for earnings inequality despite HK acquisition and learning
   - ▷ But conventional measures of sorting typically *underestimate* its impact on earnings inequality
   - ▷ Because they conflate compensating differential in wages with firm-worker match effects

Today: simulation-based results and direction of LEHD results (under disclosure)

# What We Do: Two Sets of Results

1. Develop semiparametric identification arguments for matching models with HK acquisition and learning
    - ▷ Represent wage distribution as a mixture over unobservables
    - ▷ Establish identification of this wage mixture under nonparametric restrictions
    - ▷ Recover primitives by integrating mixture-based approach w/ quantile methods for standard Roy models
2. Use it to reevaluate impact of worker-firm sorting on U.S. earnings inequality using LEHD data
    - ▷ Findings: sorting matters for earnings inequality despite HK acquisition and learning
    - ▷ But conventional measures of sorting typically *underestimate* its impact on earnings inequality
    - ▷ Because they conflate compensating differential in wages with firm-worker match effects

Today: simulation-based results and direction of LEHD results (under disclosure)

# What We Do: Two Sets of Results

1. Develop semiparametric identification arguments for matching models with HK acquisition and learning
   - ▷ Represent wage distribution as a mixture over unobservables
   - ▷ Establish identification of this wage mixture under nonparametric restrictions
   - ▷ Recover primitives by integrating mixture-based approach w/ quantile methods for standard Roy models
2. Use it to reevaluate impact of worker-firm sorting on U.S. earnings inequality using LEHD data
   - ▷ Findings: sorting matters for earnings inequality despite HK acquisition and learning
   - ▷ But conventional measures of sorting typically *underestimate* its impact on earnings inequality
   - ▷ Because they conflate compensating differential in wages with firm-worker match effects

Today: simulation-based results and direction of LEHD results (under disclosure)

# What We Do: Two Sets of Results

1. Develop semiparametric identification arguments for matching models with HK acquisition and learning
   - ▷ Represent wage distribution as a mixture over unobservables
   - ▷ Establish identification of this wage mixture under nonparametric restrictions
   - ▷ Recover primitives by integrating mixture-based approach w/ quantile methods for standard Roy models
2. Use it to reevaluate impact of worker-firm sorting on U.S. earnings inequality using LEHD data
   - ▷ Findings: sorting matters for earnings inequality despite HK acquisition and learning
   - ▷ But conventional measures of sorting typically *underestimate* its impact on earnings inequality
   - ▷ Because they conflate compensating differential in wages with firm-worker match effects

Today: simulation-based results and direction of LEHD results (under disclosure)

# A Dynamic Matching Model of the Labor Market

# Preliminaries

- *Dynamic* matching model w/ *imperfect* firm competition under uncertainty and *learning* about worker ability
- Firms *heterogeneous* in their technology of output, HK and information production
- Workers *heterogenous* in HK (observed) and "ability" (unobserved but learnt over time)
- This general framework *nests many* existing ones: models of
  - ▷ wage growth and inequality (Becker, 1975; Mincer 1958, 1974; Ben-Porath, 1967; etc.)
  - ▷ job choice with exogenous wages and no learning (à la Keane & Wolpin, 1997)
  - ▷ ... with endogenous wages, identical firms, learning (Farber & Gibbons, 1996; Gibbons & Waldman, 1999a,b)

Next: firm-workers-human capital-output-information structure

# Preliminaries

- *Dynamic* matching model w/ *imperfect* firm competition under uncertainty and *learning* about worker ability

- Firms *heterogeneous* in their technology of output, HK and information production

- Workers *heterogenous* in HK (observed) and "ability" (unobserved but learnt over time)

- This general framework *nests many* existing ones: models of
  - ▷ wage growth and inequality (Becker, 1975; Mincer 1958, 1974; Ben-Porath, 1967; etc.)
  - ▷ job choice with exogenous wages and no learning (à la Keane & Wolpin, 1997)
  - ▷ ... with endogenous wages, identical firms, learning (Farber & Gibbons, 1996; Gibbons & Waldman, 1999a,b)

Next: firm-workers-human capital-output-information structure

# Firms

- Finitely many firms produce an homogeneous good sold in perfectly competitive market at price of 1
- Production in each firm $d \in \mathcal{D}$ is governed by CRS technology in workers' labor (described in a few slides)
- Firms Bertrand-compete for workers by offering wages each period for their employment during the period
- Can easily be extended to multi-job firms where offers include both a wage and a job

# Firms

- Finitely many firms produce an homogeneous good sold in perfectly competitive market at price of 1
- Production in each firm $d \in \mathcal{D}$ is governed by CRS technology in workers' labor (described in a few slides)
- Firms Bertrand-compete for workers by offering wages each period for their employment during the period
- Can easily be extended to multi-job firms where offers include both a wage and a job

# Firms

- Finitely many firms produce an homogeneous good sold in perfectly competitive market at price of 1
- Production in each firm $d \in \mathcal{D}$ is governed by CRS technology in workers' labor (described in a few slides)
- Firms Bertrand-compete for workers by offering wages each period for their employment during the period
- Can easily be extended to multi-job firms where offers include both a wage and a job

# Firms

- Finitely many firms produce an homogeneous good sold in perfectly competitive market at price of 1
- Production in each firm $d \in \mathcal{D}$ is governed by CRS technology in workers' labor (described in a few slides)
- Firms Bertrand-compete for workers by offering wages each period for their employment during the period
- Can easily be extended to multi-job firms where offers include both a wage and a job

# Workers: Characteristics at Entry into Labor Market

- $H_{n,1}$: gender, race and initial HK (e.g. education) observed by workers, firms and econometrician
- Other skills, unobserved by the econometrician, present from birth or developed pre-market entry
  - ▷ $e_n$: *efficiency* observed by workers and firms
  - ▷ $\theta_n$: *ability* gradually and symmetrically *learnt* by workers and firms over time
- $(e_n, \theta_n)$ are *general* traits that potentially influence performance across *all* jobs (in given market/occupation)
- In the econometric part, $e_n$ is assumed discrete
- $\theta_n \in \Theta := \{\bar{\theta}, \underline{\theta}\}$: simplifies model description but can be generalized to continuous/multidimensional

# Workers: Characteristics at Entry into Labor Market

- $H_{n,1}$: gender, race and initial HK (e.g. education) observed by workers, firms and econometrician
- Other skills, unobserved by the econometrician, present from birth or developed pre-market entry
  - ▷ $e_n$: *efficiency* observed by workers and firms
  - ▷ $\theta_n$: *ability* gradually and symmetrically *learnt* by workers and firms over time
- $(e_n, \theta_n)$ are *general* traits that potentially influence performance across *all* jobs (in given market/occupation)

- In the econometric part, $e_n$ is assumed discrete
- $\theta_n \in \Theta := \{\bar{\theta}, \underline{\theta}\}$: simplifies model description but can be generalized to continuous/multidimensional

# Workers: Characteristics at Entry into Labor Market

- $H_{n,1}$: gender, race and initial HK (e.g. education) observed by workers, firms and econometrician
- Other skills, unobserved by the econometrician, present from birth or developed pre-market entry
  - $e_n$: *efficiency* observed by workers and firms
  - $\theta_n$: *ability* gradually and symmetrically *learnt* by workers and firms over time
- $(e_n, \theta_n)$ are *general* traits that potentially influence performance across *all* jobs (in given market/occupation)
- In the econometric part, $e_n$ is assumed discrete
- $\theta_n \in \Theta := \{\bar{\theta}, \underline{\theta}\}$: simplifies model description but can be generalized to continuous/multidimensional

# Workers: Characteristics at Entry into Labor Market

- $H_{n,1}$: gender, race and initial HK (e.g. education) observed by workers, firms and econometrician
- Other skills, unobserved by the econometrician, present from birth or developed pre-market entry
  - ▷ $e_n$: *efficiency* observed by workers and firms
  - ▷ $\theta_n$: *ability* gradually and symmetrically *learnt* by workers and firms over time
- $(e_n, \theta_n)$ are *general* traits that potentially influence performance across *all* jobs (in given market/occupation)
- In the econometric part, $e_n$ is assumed discrete
- $\theta_n \in \Theta := \{\bar{\theta}, \underline{\theta}\}$: simplifies model description but can be generalized to continuous/multidimensional

# Workers: Characteristics at Entry into Labor Market

- $H_{n,1}$: gender, race and initial HK (e.g. education) observed by workers, firms and econometrician
- Other skills, unobserved by the econometrician, present from birth or developed pre-market entry
  - ▷ $e_n$: *efficiency* observed by workers and firms
  - ▷ $\theta_n$: *ability* gradually and symmetrically *learnt* by workers and firms over time
- $(e_n, \theta_n)$ are *general* traits that potentially influence performance across *all* jobs (in given market/occupation)

- In the econometric part, $e_n$ is assumed discrete
- $\theta_n \in \Theta := \{\bar{\theta}, \underline{\theta}\}$: simplifies model description but can be generalized to continuous/multidimensional

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d,e)$ at *end* of $t$

$$H_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d,e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d,e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d,e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d,e)$ via *distribution* of $a_{n,t}(d,e)$
- $a_{n,t}(d,e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d, e)$ at *end* of $t$

$$H_{n,t}(d, e) = a_{n,t}(d, e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d, e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d, e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d, e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d, e)$ via *distribution* of $a_{n,t}(d, e)$
- $a_{n,t}(d, e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d,e)$ at *end* of $t$

$$H_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d,e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d,e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d,e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d,e)$ via *distribution* of $a_{n,t}(d,e)$
- $a_{n,t}(d,e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d,e)$ at *end* of $t$

$$H_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d,e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d,e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d,e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d,e)$ via *distribution* of $a_{n,t}(d,e)$
- $a_{n,t}(d,e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d, e)$ at *end* of $t$

$$H_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d,e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d,e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d,e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d,e)$ via *distribution* of $a_{n,t}(d,e)$
- $a_{n,t}(d,e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d, e)$ at *end* of $t$

$$H_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d,e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d,e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d,e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d,e)$ via *distribution* of $a_{n,t}(d,e)$
- $a_{n,t}(d,e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d, e)$ at *end* of $t$

$$H_{n,t}(d, e) = a_{n,t}(d, e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d, e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d, e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d, e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d, e)$ via *distribution* of $a_{n,t}(d, e)$
- $a_{n,t}(d, e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d,e)$ at *end* of $t$

$$H_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d,e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d,e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d,e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d,e)$ via *distribution* of $a_{n,t}(d,e)$
- $a_{n,t}(d,e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d,e)$ at *end* of $t$

$$H_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d,e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d,e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d,e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d,e)$ via *distribution* of $a_{n,t}(d,e)$
- $a_{n,t}(d,e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d,e)$ at *end* of $t$

$$H_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d,e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d,e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d,e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d,e)$ via *distribution* of $a_{n,t}(d,e)$
- $a_{n,t}(d,e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d,e)$ at *end* of $t$

$$H_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d,e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d,e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d,e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d,e)$ via *distribution* of $a_{n,t}(d,e)$
- $a_{n,t}(d,e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Workers: Human Capital Accumulation

- Worker $n$ accumulates HK over time depending on initial characteristics $(H_{n,1}, e_n, \theta_n)$ and job history $D_n^t$
  - ▷ Notation: $D_{n,t}$ is worker $n$'s job choice at time $t$ and $D_n^t := (D_{n,1}, \ldots, D_{n,t})$ is job history
  - ▷ Note (next): we allow entire job history to potentially affect acquired HK
- Worker $n$ with efficiency $e_n = e$ employed at firm $d$ in period $t$ has HK $H_{n,t}(d,e)$ at *end* of $t$

$$H_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa(D_n^{t-1})) + \epsilon_{n,t}(d,e)$$

  - ▷ Labor input: match-specific function $\ell_{d,e}(\cdot)$ of $H_{n,1}$, $\kappa_{n,t} := \kappa(D_n^{t-1})$ and productivity shock $\epsilon_{n,t}(d,e)$
  - ▷ TFP: match-specific random variable $a_{n,t}(d,e)$ with *distribution* governed by $d, e, H_{n,1}$ and $\theta_n$
- So unknown ability $\theta_n$ influences $H_{n,t}(d,e)$ via *distribution* of $a_{n,t}(d,e)$
- $a_{n,t}(d,e)$ is a *noisy* measure of $\theta_n$ *not* a deterministic function (for learning not to be trivial)

# Output Technology

- With labor supply normalized to 1, output produced is equal to total HK: $y_{n,t}(d,e) = H_{n,t}(d,e)$

- Worker $n$ with efficiency $e_n = e$ employed at $d$ in period $t$ produces output $y_{n,t}(d,e)$ at *end* of $t$

$$y_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa_{n,t}) + \epsilon_{n,t}(d,e)$$

- Firms differ in their output/HK technology ($\ell_{d,e}$ and distribution of $a_{n,t}(d,e)$ and $\epsilon_{n,t}(d,e)$ depend on $d$)

- Next: discuss how firms differ in their technology of *information* generation

# Output Technology

- With labor supply normalized to 1, output produced is equal to total HK: $y_{n,t}(d,e) = H_{n,t}(d,e)$
- Worker $n$ with efficiency $e_n = e$ employed at $d$ in period $t$ produces output $y_{n,t}(d,e)$ at *end* of $t$

$$y_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa_{n,t}) + \epsilon_{n,t}(d,e)$$

- Firms differ in their output/HK technology ($\ell_{d,e}$ and distribution of $a_{n,t}(d,e)$ and $\epsilon_{n,t}(d,e)$ depend on $d$)
- Next: discuss how firms differ in their technology of *information* generation

# Output Technology

- With labor supply normalized to 1, output produced is equal to total HK: $y_{n,t}(d,e) = H_{n,t}(d,e)$

- Worker $n$ with efficiency $e_n = e$ employed at $d$ in period $t$ produces output $y_{n,t}(d,e)$ at *end* of $t$

$$y_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa_{n,t}) + \epsilon_{n,t}(d,e)$$

- Firms differ in their output/HK technology ($\ell_{d,e}$ and distribution of $a_{n,t}(d,e)$ and $\epsilon_{n,t}(d,e)$ depend on $d$)

- Next: discuss how firms differ in their technology of *information* generation

# Output Technology

- With labor supply normalized to 1, output produced is equal to total HK: $y_{n,t}(d,e) = H_{n,t}(d,e)$

- Worker $n$ with efficiency $e_n = e$ employed at $d$ in period $t$ produces output $y_{n,t}(d,e)$ at *end* of $t$

$$y_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa_{n,t}) + \epsilon_{n,t}(d,e)$$

- Firms differ in their output/HK technology ($\ell_{d,e}$ and distribution of $a_{n,t}(d,e)$ and $\epsilon_{n,t}(d,e)$ depend on $d$)

- Next: discuss how firms differ in their technology of *information* generation

# Information

- Recall output produced by worker $n$ with efficiency $e_n = e$ employed at $d$ in period $t$ at *end* of $t$ is

$$y_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa_{n,t}) + \epsilon_{n,t}(d,e)$$

- *Beginning* of $t$ (before decisions): workers and firms know $\{\ell_{d,e}(H_{n,1}, \kappa_{n,t})\}_d$ and observe $\{\epsilon_{n,t}(d,e)\}_d$
- *End* of $t$ (after decisions): workers and firms observe $a_{n,t}(d,e)$ at employing firm $d$
- In particular, workers and firms make decisions based on *expectation* about $\{a_{n,t}(d,e)\}_d$
- Key: *distribution* of $\{a_{n,t}(d,e)\}_d$ depends on unknown ability $\theta_n$
- Therefore, expectation about $\{a_{n,t}(d,e)\}_d$ depends on beliefs about $\theta_n$

# Information

- Recall output produced by worker $n$ with efficiency $e_n = e$ employed at $d$ in period $t$ at *end* of $t$ is

$$y_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa_{n,t}) + \epsilon_{n,t}(d,e)$$

- *Beginning* of $t$ (before decisions): workers and firms know $\{\ell_{d,e}(H_{n,1}, \kappa_{n,t})\}_d$ and observe $\{\epsilon_{n,t}(d,e)\}_d$
- *End* of $t$ (after decisions): workers and firms observe $a_{n,t}(d,e)$ at employing firm $d$
- In particular, workers and firms make decisions based on *expectation* about $\{a_{n,t}(d,e)\}_d$
- Key: *distribution* of $\{a_{n,t}(d,e)\}_d$ depends on unknown ability $\theta_n$
- Therefore, expectation about $\{a_{n,t}(d,e)\}_d$ depends on beliefs about $\theta_n$

# Information

- Recall output produced by worker $n$ with efficiency $e_n = e$ employed at $d$ in period $t$ at *end* of $t$ is

$$y_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa_{n,t}) + \epsilon_{n,t}(d,e)$$

- *Beginning* of $t$ (before decisions): workers and firms know $\{\ell_{d,e}(H_{n,1}, \kappa_{n,t})\}_d$ and observe $\{\epsilon_{n,t}(d,e)\}_d$
- *End* of $t$ (after decisions): workers and firms observe $a_{n,t}(d,e)$ at employing firm $d$
- In particular, workers and firms make decisions based on *expectation* about $\{a_{n,t}(d,e)\}_d$
- Key: *distribution* of $\{a_{n,t}(d,e)\}_d$ depends on unknown ability $\theta_n$
- Therefore, expectation about $\{a_{n,t}(d,e)\}_d$ depends on beliefs about $\theta_n$

# Information

- Recall output produced by worker $n$ with efficiency $e_n = e$ employed at $d$ in period $t$ at *end* of $t$ is

$$y_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa_{n,t}) + \epsilon_{n,t}(d,e)$$

- *Beginning* of $t$ (before decisions): workers and firms know $\{\ell_{d,e}(H_{n,1}, \kappa_{n,t})\}_d$ and observe $\{\epsilon_{n,t}(d,e)\}_d$
- *End* of $t$ (after decisions): workers and firms observe $a_{n,t}(d,e)$ at employing firm $d$
- In particular, workers and firms make decisions based on *expectation* about $\{a_{n,t}(d,e)\}_d$
- Key: *distribution* of $\{a_{n,t}(d,e)\}_d$ depends on unknown ability $\theta_n$
- Therefore, expectation about $\{a_{n,t}(d,e)\}_d$ depends on beliefs about $\theta_n$

# Information

- Recall output produced by worker $n$ with efficiency $e_n = e$ employed at $d$ in period $t$ at *end* of $t$ is

$$y_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa_{n,t}) + \epsilon_{n,t}(d,e)$$

- *Beginning* of $t$ (before decisions): workers and firms know $\{\ell_{d,e}(H_{n,1}, \kappa_{n,t})\}_d$ and observe $\{\epsilon_{n,t}(d,e)\}_d$
- *End* of $t$ (after decisions): workers and firms observe $a_{n,t}(d,e)$ at employing firm $d$
- In particular, workers and firms make decisions based on *expectation* about $\{a_{n,t}(d,e)\}_d$
- Key: *distribution* of $\{a_{n,t}(d,e)\}_d$ depends on unknown ability $\theta_n$
- Therefore, expectation about $\{a_{n,t}(d,e)\}_d$ depends on beliefs about $\theta_n$

# Information

- Recall output produced by worker $n$ with efficiency $e_n = e$ employed at $d$ in period $t$ at *end* of $t$ is

$$y_{n,t}(d,e) = a_{n,t}(d,e) + \ell_{d,e}(H_{n,1}, \kappa_{n,t}) + \epsilon_{n,t}(d,e)$$

- *Beginning* of $t$ (before decisions): workers and firms know $\{\ell_{d,e}(H_{n,1}, \kappa_{n,t})\}_d$ and observe $\{\epsilon_{n,t}(d,e)\}_d$
- *End* of $t$ (after decisions): workers and firms observe $a_{n,t}(d,e)$ at employing firm $d$
- In particular, workers and firms make decisions based on *expectation* about $\{a_{n,t}(d,e)\}_d$
- Key: *distribution* of $\{a_{n,t}(d,e)\}_d$ depends on unknown ability $\theta_n$
- Therefore, expectation about $\{a_{n,t}(d,e)\}_d$ depends on beliefs about $\theta_n$

# Learning about Ability

- Workers and firms symmetrically learn about $\theta_n$ through Bayesian updating process
- This process is based on common observations of $a_{n,t}(d,e)$ at end of period $t$ at employing firm $d$
- How does it work? Workers and firms have prior belief about $\theta_n = \bar{\theta}$ at *beginning* of $t = 1$

$$P_{n,1} = p_1(h,e) := \Pr(\theta_n = \bar{\theta} \mid H_{n,1} = h, e_n = e)$$

- At *end* of $t \geq 1$ they observe $a_{n,t}(d,e)$ and so $y_{n,t}(d,e)$ at employing firm $d$
- At *beginning* of $t+1$ they update beliefs about $\theta_n$ using signal $a_{n,t}(d,e)$ and Bayes' rule to obtain $P_{n,t+1}$

# Learning about Ability

- Workers and firms symmetrically learn about $\theta_n$ through Bayesian updating process
- This process is based on common observations of $a_{n,t}(d,e)$ at end of period $t$ at employing firm $d$
- How does it work? Workers and firms have prior belief about $\theta_n = \bar{\theta}$ at *beginning* of $t = 1$

$$P_{n,1} = p_1(h,e) := \Pr(\theta_n = \bar{\theta} \mid H_{n,1} = h, e_n = e)$$

- At *end* of $t \geq 1$ they observe $a_{n,t}(d,e)$ and so $y_{n,t}(d,e)$ at employing firm $d$
- At *beginning* of $t+1$ they update beliefs about $\theta_n$ using signal $a_{n,t}(d,e)$ and Bayes' rule to obtain $P_{n,t+1}$

# Learning about Ability

- Workers and firms symmetrically learn about $\theta_n$ through Bayesian updating process
- This process is based on common observations of $a_{n,t}(d,e)$ at end of period $t$ at employing firm $d$
- How does it work? Workers and firms have prior belief about $\theta_n = \bar{\theta}$ at *beginning* of $t = 1$

$$P_{n,1} = p_1(h,e) := \Pr(\theta_n = \bar{\theta} \mid H_{n,1} = h, e_n = e)$$

- At *end* of $t \geq 1$ they observe $a_{n,t}(d,e)$ and so $y_{n,t}(d,e)$ at employing firm $d$
- At *beginning* of $t+1$ they update beliefs about $\theta_n$ using signal $a_{n,t}(d,e)$ and Bayes' rule to obtain $P_{n,t+1}$

# Learning about Ability

- Workers and firms symmetrically learn about $\theta_n$ through Bayesian updating process
- This process is based on common observations of $a_{n,t}(d,e)$ at end of period $t$ at employing firm $d$
- How does it work? Workers and firms have prior belief about $\theta_n = \bar{\theta}$ at *beginning* of $t=1$

$$P_{n,1} = p_1(h,e) := \Pr(\theta_n = \bar{\theta} \mid H_{n,1} = h, e_n = e)$$

- At *end* of $t \geq 1$ they observe $a_{n,t}(d,e)$ and so $y_{n,t}(d,e)$ at employing firm $d$
- At *beginning* of $t+1$ they update beliefs about $\theta_n$ using signal $a_{n,t}(d,e)$ and Bayes' rule to obtain $P_{n,t+1}$

# Learning about Ability

- Workers and firms symmetrically learn about $\theta_n$ through Bayesian updating process

- This process is based on common observations of $a_{n,t}(d,e)$ at end of period $t$ at employing firm $d$

- How does it work? Workers and firms have prior belief about $\theta_n = \bar{\theta}$ at *beginning* of $t = 1$

$$P_{n,1} = p_1(h,e) := \Pr(\theta_n = \bar{\theta} \mid H_{n,1} = h, e_n = e)$$

- At *end* of $t \geq 1$ they observe $a_{n,t}(d,e)$ and so $y_{n,t}(d,e)$ at employing firm $d$

- At *beginning* of $t+1$ they update beliefs about $\theta_n$ using signal $a_{n,t}(d,e)$ and Bayes' rule to obtain $P_{n,t+1}$

# Expected Output

- Given these beliefs, workers and firms make decisions based on *expected output* in period $t+1$
- Expected output of worker $n$ in period $t+1$ when employed at firm $d$ (before $a_{n,t+1}(d,e)$ is realized)

$$\mathbb{E}\big(y_{n,t+1}(d,e) \mid P_{n,t+1}=p, e_n=e, H_{n,1}=h, \kappa_{n,t+1}=\kappa, \epsilon_{n,t+1}\big) = y(d,s) + \epsilon_{n,t+1}(d,e)$$

- Conditional on all that is known by workers and firms at beginning of $t+1$
- $y(s,d)$ is expected value of $y_{n,t+1}(d,e)$ given $s_{n,t+1} := (P_{n,t+1}, e_n, H_{n,1}, \kappa_{n,t+1})$ w/ realization $s := (p,e,h,\kappa)$
- $\epsilon_{n,t+1} := (\epsilon_{n,t+1}(d,e) : d \in \mathcal{D}, e \in \mathcal{E})$ is vector of shocks
- Beginning-of-period state variables: $(s_{n,t+1}, \epsilon_{n,t+1})$
- $y(d,s)$ is primitive of interest as it captures heterogeneous output/HK technology of firms

Next: equilibrium definition

# Expected Output

- Given these beliefs, workers and firms make decisions based on *expected output* in period $t+1$
- Expected output of worker $n$ in period $t+1$ when employed at firm $d$ (before $a_{n,t+1}(d,e)$ is realized)

$$\mathbb{E}\big(y_{n,t+1}(d,e) \mid P_{n,t+1} = p, e_n = e, H_{n,1} = h, \kappa_{n,t+1} = \kappa, \epsilon_{n,t+1}\big) = y(d,s) + \epsilon_{n,t+1}(d,e)$$

- Conditional on all that is known by workers and firms at beginning of $t+1$
- $y(s,d)$ is expected value of $y_{n,t+1}(d,e)$ given $s_{n,t+1} := (P_{n,t+1}, e_n, H_{n,1}, \kappa_{n,t+1})$ w/ realization $s := (p,e,h,\kappa)$
- $\epsilon_{n,t+1} := (\epsilon_{n,t+1}(d,e) : d \in \mathcal{D}, e \in \mathcal{E})$ is vector of shocks
- Beginning-of-period state variables: $(s_{n,t+1}, \epsilon_{n,t+1})$
- $y(d,s)$ is primitive of interest as it captures heterogeneous output/HK technology of firms

Next: equilibrium definition

# Expected Output

- Given these beliefs, workers and firms make decisions based on *expected output* in period $t+1$
- Expected output of worker $n$ in period $t+1$ when employed at firm $d$ (before $a_{n,t+1}(d,e)$ is realized)

$$\mathbb{E}\big(y_{n,t+1}(d,e) \mid P_{n,t+1} = p, e_n = e, H_{n,1} = h, \kappa_{n,t+1} = \kappa, \epsilon_{n,t+1}\big) = y(d,s) + \epsilon_{n,t+1}(d,e)$$

- Conditional on all that is known by workers and firms at beginning of $t+1$
- $y(s,d)$ is expected value of $y_{n,t+1}(d,e)$ given $s_{n,t+1} := (P_{n,t+1}, e_n, H_{n,1}, \kappa_{n,t+1})$ w/ realization $s := (p,e,h,\kappa)$
- $\epsilon_{n,t+1} := (\epsilon_{n,t+1}(d,e) : d \in \mathcal{D}, e \in \mathcal{E})$ is vector of shocks
- Beginning-of-period state variables: $(s_{n,t+1}, \epsilon_{n,t+1})$
- $y(d,s)$ is primitive of interest as it captures heterogeneous output/HK technology of firms

Next: equilibrium definition

# Expected Output

- Given these beliefs, workers and firms make decisions based on *expected output* in period $t+1$
- Expected output of worker $n$ in period $t+1$ when employed at firm $d$ (before $a_{n,t+1}(d,e)$ is realized)

$$\mathbb{E}\big(y_{n,t+1}(d,e) \mid P_{n,t+1} = p, e_n = e, H_{n,1} = h, \kappa_{n,t+1} = \kappa, \epsilon_{n,t+1}\big) = y(d,s) + \epsilon_{n,t+1}(d,e)$$

- Conditional on all that is known by workers and firms at beginning of $t+1$
- $y(s,d)$ is expected value of $y_{n,t+1}(d,e)$ given $s_{n,t+1} := (P_{n,t+1}, e_n, H_{n,1}, \kappa_{n,t+1})$ w/ realization $s := (p,e,h,\kappa)$
- $\epsilon_{n,t+1} := (\epsilon_{n,t+1}(d,e) : d \in \mathcal{D}, e \in \mathcal{E})$ is vector of shocks
- Beginning-of-period state variables: $(s_{n,t+1}, \epsilon_{n,t+1})$
- $y(d,s)$ is primitive of interest as it captures heterogeneous output/HK technology of firms

Next: equilibrium definition

# Expected Output

- Given these beliefs, workers and firms make decisions based on *expected output* in period $t+1$
- Expected output of worker $n$ in period $t+1$ when employed at firm $d$ (before $a_{n,t+1}(d,e)$ is realized)

$$\mathbb{E}\big(y_{n,t+1}(d,e) \mid P_{n,t+1}=p, e_n=e, H_{n,1}=h, \kappa_{n,t+1}=\kappa, \epsilon_{n,t+1}\big) = y(d,s) + \epsilon_{n,t+1}(d,e)$$

- Conditional on all that is known by workers and firms at beginning of $t+1$
- $y(s,d)$ is expected value of $y_{n,t+1}(d,e)$ given $s_{n,t+1} := (P_{n,t+1}, e_n, H_{n,1}, \kappa_{n,t+1})$ w/ realization $s := (p,e,h,\kappa)$
- $\epsilon_{n,t+1} := (\epsilon_{n,t+1}(d,e) : d \in \mathcal{D}, e \in \mathcal{E})$ is vector of shocks
- Beginning-of-period state variables: $(s_{n,t+1}, \epsilon_{n,t+1})$
- $y(d,s)$ is primitive of interest as it captures heterogeneous output/HK technology of firms

Next: equilibrium definition

# Expected Output

- Given these beliefs, workers and firms make decisions based on *expected output* in period $t+1$
- Expected output of worker $n$ in period $t+1$ when employed at firm $d$ (before $a_{n,t+1}(d,e)$ is realized)

$$\mathbb{E}\left(y_{n,t+1}(d,e) \mid P_{n,t+1} = p, e_n = e, H_{n,1} = h, \kappa_{n,t+1} = \kappa, \epsilon_{n,t+1}\right) = y(d,s) + \epsilon_{n,t+1}(d,e)$$

- Conditional on all that is known by workers and firms at beginning of $t+1$
- $y(s,d)$ is expected value of $y_{n,t+1}(d,e)$ given $s_{n,t+1} := (P_{n,t+1}, e_n, H_{n,1}, \kappa_{n,t+1})$ w/ realization $s := (p,e,h,\kappa)$
- $\epsilon_{n,t+1} := (\epsilon_{n,t+1}(d,e) : d \in \mathcal{D}, e \in \mathcal{E})$ is vector of shocks
- Beginning-of-period state variables: $(s_{n,t+1}, \epsilon_{n,t+1})$
- $y(d,s)$ is primitive of interest as it captures heterogeneous output/HK technology of firms

Next: equilibrium definition

# Expected Output

- Given these beliefs, workers and firms make decisions based on *expected output* in period $t+1$
- Expected output of worker $n$ in period $t+1$ when employed at firm $d$ (before $a_{n,t+1}(d,e)$ is realized)

$$\mathbb{E}\big(y_{n,t+1}(d,e) \mid P_{n,t+1}=p, e_n=e, H_{n,1}=h, \kappa_{n,t+1}=\kappa, \epsilon_{n,t+1}\big) = y(d,s) + \epsilon_{n,t+1}(d,e)$$

- Conditional on all that is known by workers and firms at beginning of $t+1$
- $y(s,d)$ is expected value of $y_{n,t+1}(d,e)$ given $s_{n,t+1} := (P_{n,t+1}, e_n, H_{n,1}, \kappa_{n,t+1})$ w/ realization $s := (p,e,h,\kappa)$
- $\epsilon_{n,t+1} := (\epsilon_{n,t+1}(d,e) : d \in \mathcal{D}, e \in \mathcal{E})$ is vector of shocks
- Beginning-of-period state variables: $(s_{n,t+1}, \epsilon_{n,t+1})$
- $y(d,s)$ is primitive of interest as it captures heterogeneous output/HK technology of firms

Next: equilibrium definition

# Equilibrium: Robust MPE

- Given absence of complementarities, we can focus on competition of all firms for one worker at a time
- A robust MPE consists of wage and acceptance strategies together with belief process such that
  - ▷ The worker maximizes the (expected present discounted) value of wages
  - ▷ Each firm maximizes the (expected present discounted) value of profits
  - ▷ Beliefs are updated according to Bayes' rule
  - ▷ Non-employing firms indifferent between employing and not the worker (guarantees *wages are unique*)
- So equilibrium exists and is unique (and efficient) (Bergemann & Valimaki, 1996) [Bellman] [Refinement]
  - ▷ Can be inefficient in the multi-job case (Pastorino, 2024)

# Equilibrium: Robust MPE

- Given absence of complementarities, we can focus on competition of all firms for one worker at a time
- A robust MPE consists of wage and acceptance strategies together with belief process such that
  - ▷ The worker maximizes the (expected present discounted) value of wages
  - ▷ Each firm maximizes the (expected present discounted) value of profits
  - ▷ Beliefs are updated according to Bayes' rule
  - ▷ Non-employing firms indifferent between employing and not the worker (guarantees *wages are unique*)
- So equilibrium exists and is unique (and efficient) (Bergemann & Valimaki, 1996) [Bellman] [Refinement]
  - ▷ Can be inefficient in the multi-job case (Pastorino, 2024)

# Equilibrium: Robust MPE

- Given absence of complementarities, we can focus on competition of all firms for one worker at a time
- A robust MPE consists of wage and acceptance strategies together with belief process such that
  - ▷ The worker maximizes the (expected present discounted) value of wages
  - ▷ Each firm maximizes the (expected present discounted) value of profits
  - ▷ Beliefs are updated according to Bayes' rule
  - ▷ Non-employing firms indifferent between employing and not the worker (guarantees *wages are unique*)
- So equilibrium exists and is unique (and efficient) (Bergemann & Valimaki, 1996) [Bellman] [Refinement]
  - ▷ Can be inefficient in the multi-job case (Pastorino, 2024)

# Equilibrium: Robust MPE

- Given absence of complementarities, we can focus on competition of all firms for one worker at a time
- A robust MPE consists of wage and acceptance strategies together with belief process such that
  - ▷ The worker maximizes the (expected present discounted) value of wages
  - ▷ Each firm maximizes the (expected present discounted) value of profits
  - ▷ Beliefs are updated according to Bayes' rule
  - ▷ Non-employing firms indifferent between employing and not the worker (guarantees *wages are unique*)
- So equilibrium exists and is unique (and efficient) (Bergemann & Valimaki, 1996) [Bellman] [Refinement]
  - ▷ Can be inefficient in the multi-job case (Pastorino, 2024)

# Equilibrium Wage: Overview

- The equilibrium wage of worker employed at firm $d$ is the sum of

    1. *Expected output* at the *2nd-best* firm $d'$ (from worker point of view: in terms of EPDV of wages)

    2. *Compensating differential*: wage premium or wage discount over worker $n$'s expected output

- This second component arises as equilibrium leads to *second-price auction-like pricing* mechanism

- Why? Turn to build intuition starting from static case

# Equilibrium Wage: Overview

- The equilibrium wage of worker employed at firm $d$ is the sum of

  1. *Expected output* at the *2nd-best* firm $d'$ (from worker point of view: in terms of EPDV of wages)

  2. *Compensating differential*: wage premium or wage discount over worker $n$'s expected output

- This second component arises as equilibrium leads to *second-price auction-like pricing* mechanism

- Why? Turn to build intuition starting from static case

# Equilibrium Wage: Overview

- The equilibrium wage of worker employed at firm $d$ is the sum of

    1. *Expected output* at the *2nd-best* firm $d'$ (from worker point of view: in terms of EPDV of wages)

    2. *Compensating differential*: wage premium or wage discount over worker $n$'s expected output

- This second component arises as equilibrium leads to *second-price auction-like pricing* mechanism

- Why? Turn to build intuition starting from static case

# Equilibrium Wage: Overview

- The equilibrium wage of worker employed at firm $d$ is the sum of

  1. *Expected output* at the *2nd-best* firm $d'$ (from worker point of view: in terms of EPDV of wages)

  2. *Compensating differential*: wage premium or wage discount over worker $n$'s expected output

- This second component arises as equilibrium leads to *second-price auction-like pricing* mechanism

- Why? Turn to build intuition starting from static case

# Equilibrium Wage: Overview

- The equilibrium wage of worker employed at firm $d$ is the sum of

  1. *Expected output* at the *2nd-best* firm $d'$ (from worker point of view: in terms of EPDV of wages)

  2. *Compensating differential*: wage premium or wage discount over worker $n$'s expected output

- This second component arises as equilibrium leads to *second-price auction-like pricing* mechanism

- Why? Turn to build intuition starting from static case

# Equilibrium Wage: Static Case with Two Firms

- Under static Bertrand competition among *differentiated* firms *selling* a common good
  - ▷ Both firms offer the same *price*
  - ▷ The high-productivity (low-cost) firm sells at price equal to cost of the low-productivity (high-cost) firm
- Under static version of our model in which two *differentiated* firms *buy* the services of a worker
  - ▷ Both firms offer the same *wage*
  - ▷ Worker's wage equals the expected output worker would produce if hired by the low-productivity firm
- Thus, the worker is indifferent between the two firms

# Equilibrium Wage: Static Case with Two Firms

- Under static Bertrand competition among *differentiated* firms *selling* a common good
  - ▷ Both firms offer the same *price*
  - ▷ The high-productivity (low-cost) firm sells at price equal to cost of the low-productivity (high-cost) firm

- Under static version of our model in which two *differentiated* firms *buy* the services of a worker
  - ▷ Both firms offer the same *wage*
  - ▷ Worker's wage equals the expected output worker would produce if hired by the low-productivity firm

- Thus, the worker is indifferent between the two firms

# Equilibrium Wage: Static Case with Two Firms

- Under static Bertrand competition among *differentiated* firms *selling* a common good
  - ▷ Both firms offer the same *price*
  - ▷ The high-productivity (low-cost) firm sells at price equal to cost of the low-productivity (high-cost) firm
- Under static version of our model in which two *differentiated* firms *buy* the services of a worker
  - ▷ Both firms offer the same *wage*
  - ▷ Worker's wage equals the expected output worker would produce if hired by the low-productivity firm
- Thus, the worker is indifferent between the two firms

# Equilibrium Wage: Dynamic Case

- Same intuition holds in the dynamic case (worker indifferent btw EPDV of wages) with two key differences
  - ▷ Firms differ in both output/HK and information technologies
  - ▷ HK and information acquired (or forgone) through employment lead to future higher (or lower) wages
  - ▷ Note: any such benefit or cost is *capitalized* in the paid wage
- Say, firm offering HK or info gains leading to higher future wages can pay lower wage yet still attract worker
- In equilibrium, a worker's wage equals the expected output the worker would produce if hired by competitor
- *Plus* premium/discount reflecting wage value of future HK/info if competitor's offer had been accepted
- This extra term enters the wage equation as a *compensating differential* to make the worker indifferent
- Same intuition with $>2$ firms: the two relevant firms are those offering the two highest EPDV of wages

# Equilibrium Wage: Dynamic Case

- Same intuition holds in the dynamic case (worker indifferent btw EPDV of wages) with two key differences
  - ▷ Firms differ in both output/HK and information technologies
  - ▷ HK and information acquired (or forgone) through employment lead to future higher (or lower) wages
  - ▷ Note: any such benefit or cost is *capitalized* in the paid wage
- Say, firm offering HK or info gains leading to higher future wages can pay lower wage yet still attract worker
- In equilibrium, a worker's wage equals the expected output the worker would produce if hired by competitor
- *Plus* premium/discount reflecting wage value of future HK/info if competitor's offer had been accepted
- This extra term enters the wage equation as a *compensating differential* to make the worker indifferent
- Same intuition with $> 2$ firms: the two relevant firms are those offering the two highest EPDV of wages

# Equilibrium Wage: Dynamic Case

- Same intuition holds in the dynamic case (worker indifferent btw EPDV of wages) with two key differences
  - ▷ Firms differ in both output/HK and information technologies
  - ▷ HK and information acquired (or forgone) through employment lead to future higher (or lower) wages
  - ▷ Note: any such benefit or cost is *capitalized* in the paid wage

- Say, firm offering HK or info gains leading to higher future wages can pay lower wage yet still attract worker

- In equilibrium, a worker's wage equals the expected output the worker would produce if hired by competitor

- *Plus* premium/discount reflecting wage value of future HK/info if competitor's offer had been accepted

- This extra term enters the wage equation as a *compensating differential* to make the worker indifferent

- Same intuition with $> 2$ firms: the two relevant firms are those offering the two highest EPDV of wages

# Equilibrium Wage: Dynamic Case

- Same intuition holds in the dynamic case (worker indifferent btw EPDV of wages) with two key differences
  - ▷ Firms differ in both output/HK and information technologies
  - ▷ HK and information acquired (or forgone) through employment lead to future higher (or lower) wages
  - ▷ Note: any such benefit or cost is *capitalized* in the paid wage
- Say, firm offering HK or info gains leading to higher future wages can pay lower wage yet still attract worker
- In equilibrium, a worker's wage equals the expected output the worker would produce if hired by competitor
- *Plus* premium/discount reflecting wage value of future HK/info if competitor's offer had been accepted
- This extra term enters the wage equation as a *compensating differential* to make the worker indifferent
- Same intuition with $> 2$ firms: the two relevant firms are those offering the two highest EPDV of wages

# Equilibrium Wage: Dynamic Case

- Same intuition holds in the dynamic case (worker indifferent btw EPDV of wages) with two key differences
  - ▷ Firms differ in both output/HK and information technologies
  - ▷ HK and information acquired (or forgone) through employment lead to future higher (or lower) wages
  - ▷ Note: any such benefit or cost is *capitalized* in the paid wage
- Say, firm offering HK or info gains leading to higher future wages can pay lower wage yet still attract worker
- In equilibrium, a worker's wage equals the expected output the worker would produce if hired by competitor
- *Plus* premium/discount reflecting wage value of future HK/info if competitor's offer had been accepted
- This extra term enters the wage equation as a *compensating differential* to make the worker indifferent
- Same intuition with $> 2$ firms: the two relevant firms are those offering the two highest EPDV of wages

# Equilibrium Wage Equation

- Consider equilibrium ranking of firms based on EPDV of wages offered to worker $n$ in period $t$
- Between the two firms ranked highest, 1st-best is employing firm $d$, 2nd-best is non-employing firm $d'$
- Equilibrium wage of worker $n$ with efficiency $e_n = e$ at time $t$

$$w_{n,t}(d,d',e) = \underbrace{y(d', s_{n,t}(e)) + \epsilon_{n,t}(d',e)}_{\text{expected output at 2nd-best firm } d'} + \underbrace{\Psi(d,d',s_{n,t}(e))}_{\text{compensating differential}}$$

- The compensating differential $\Psi(d,d',s_{n,t}(e))$ is difference between two value functions

$$\Psi(d,d',s_{n,t}(e)) := \overbrace{\delta[1-\eta(d',\kappa_{n,t})]\int_{\epsilon_{n,t+1}(e)} \mathbb{E}V_{d'}(s_{n,t+1}(e),\epsilon_{n,t+1}(e)|s_{n,t}(e);d')dG_e}^{\text{EPDV of match surplus of } \{d',n\} \text{ had } d' \text{ employed } n \text{ in } t \text{ (counterfactual)}}$$
$$\underbrace{-\,\delta[1-\eta(d,\kappa_{n,t})]\int_{\epsilon_{n,t+1}(e)} \mathbb{E}V_{d'}(s_{n,t+1}(e),\epsilon_{n,t+1}(e)|s_{n,t}(e);d)dG_e}_{\text{EPDV of match surplus of } \{d',n\} \text{ when } d \text{ employs } n \text{ in } t}$$

- This wage equation follows immediately from worker indifference and non-employing firm indifference

# Equilibrium Wage Equation

- Consider equilibrium ranking of firms based on EPDV of wages offered to worker $n$ in period $t$
- Between the two firms ranked highest, 1st-best is employing firm $d$, 2nd-best is non-employing firm $d'$
- Equilibrium wage of worker $n$ with efficiency $e_n = e$ at time $t$

$$w_{n,t}(d, d', e) = \underbrace{y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{expected output at 2nd-best firm } d'} + \underbrace{\Psi(d, d', s_{n,t}(e))}_{\text{compensating differential}}$$

- The compensating differential $\Psi(d, d', s_{n,t}(e))$ is difference between two value functions

$$\Psi(d, d', s_{n,t}(e)) := \overbrace{\delta[1 - \eta(d', \kappa_{n,t})] \int_{\epsilon_{n,t+1}(e)} \mathbb{E} V_{d'}(s_{n,t+1}(e), \epsilon_{n,t+1}(e) | s_{n,t}(e); d') dG_e}^{\text{EPDV of match surplus of } \{d', n\} \text{ had } d' \text{ employed } n \text{ in } t \text{ (counterfactual)}}$$
$$\underbrace{- \delta[1 - \eta(d, \kappa_{n,t})] \int_{\epsilon_{n,t+1}(e)} \mathbb{E} V_{d'}(s_{n,t+1}(e), \epsilon_{n,t+1}(e) | s_{n,t}(e); d) dG_e}_{\text{EPDV of match surplus of } \{d', n\} \text{ when } d \text{ employs } n \text{ in } t}$$

- This wage equation follows immediately from worker indifference and non-employing firm indifference

# Equilibrium Wage Equation

- Consider equilibrium ranking of firms based on EPDV of wages offered to worker $n$ in period $t$
- Between the two firms ranked highest, 1st-best is employing firm $d$, 2nd-best is non-employing firm $d'$
- Equilibrium wage of worker $n$ with efficiency $e_n = e$ at time $t$

$$w_{n,t}(d,d',e) = \underbrace{y(d', s_{n,t}(e)) + \epsilon_{n,t}(d',e)}_{\text{expected output at 2nd-best firm } d'} + \underbrace{\Psi(d,d',s_{n,t}(e))}_{\text{compensating differential}}$$

- The compensating differential $\Psi(d,d',s_{n,t}(e))$ is difference between two value functions

$$\Psi(d,d',s_{n,t}(e)) := \overbrace{\delta[1-\eta(d',\kappa_{n,t})] \int_{\epsilon_{n,t+1}(e)} \mathbb{E} V_{d'}(s_{n,t+1}(e), \epsilon_{n,t+1}(e) | s_{n,t}(e); d') dG_e}^{\text{EPDV of match surplus of } \{d',n\} \text{ had } d' \text{ employed } n \text{ in } t \text{ (counterfactual)}}$$
$$\underbrace{- \delta[1-\eta(d,\kappa_{n,t})] \int_{\epsilon_{n,t+1}(e)} \mathbb{E} V_{d'}(s_{n,t+1}(e), \epsilon_{n,t+1}(e) | s_{n,t}(e); d) dG_e}_{\text{EPDV of match surplus of } \{d',n\} \text{ when } d \text{ employs } n \text{ in } t}$$

- This wage equation follows immediately from worker indifference and non-employing firm indifference

# Equilibrium Wage Equation

- Consider equilibrium ranking of firms based on EPDV of wages offered to worker $n$ in period $t$
- Between the two firms ranked highest, 1st-best is employing firm $d$, 2nd-best is non-employing firm $d'$
- Equilibrium wage of worker $n$ with efficiency $e_n = e$ at time $t$

$$w_{n,t}(d, d', e) = \underbrace{y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{expected output at 2nd-best firm } d'} + \underbrace{\Psi(d, d', s_{n,t}(e))}_{\text{compensating differential}}$$

- The compensating differential $\Psi(d, d', s_{n,t}(e))$ is difference between two value functions

$$\begin{aligned}\Psi(d, d', s_{n,t}(e)) :=& \overbrace{\delta[1-\eta(d', \kappa_{n,t})] \int_{\epsilon_{n,t+1}(e)} \mathbb{E} V_{d'}(s_{n,t+1}(e), \epsilon_{n,t+1}(e) | s_{n,t}(e); d') dG_e}^{\text{EPDV of match surplus of } \{d', n\} \text{ had } d' \text{ employed } n \text{ in } t \text{ (counterfactual)}} \\ & \underbrace{- \delta[1-\eta(d, \kappa_{n,t})] \int_{\epsilon_{n,t+1}(e)} \mathbb{E} V_{d'}(s_{n,t+1}(e), \epsilon_{n,t+1}(e) | s_{n,t}(e); d) dG_e}_{\text{EPDV of match surplus of } \{d', n\} \text{ when } d \text{ employs } n \text{ in } t}\end{aligned}$$

- This wage equation follows immediately from worker indifference and non-employing firm indifference

# Equilibrium Wage Equation

- Consider equilibrium ranking of firms based on EPDV of wages offered to worker $n$ in period $t$
- Between the two firms ranked highest, 1st-best is employing firm $d$, 2nd-best is non-employing firm $d'$
- Equilibrium wage of worker $n$ with efficiency $e_n = e$ at time $t$

$$w_{n,t}(d, d', e) = \underbrace{y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{expected output at 2nd-best firm } d'} + \underbrace{\Psi(d, d', s_{n,t}(e))}_{\text{compensating differential}}$$

- The compensating differential $\Psi(d, d', s_{n,t}(e))$ is difference between two value functions

$$\begin{aligned}\Psi(d, d', s_{n,t}(e)) :=& \overbrace{\delta[1-\eta(d', \kappa_{n,t})] \int_{\epsilon_{n,t+1}(e)} \mathbb{E} V_{d'}(s_{n,t+1}(e), \epsilon_{n,t+1}(e) | s_{n,t}(e); d') dG_e}^{\text{EPDV of match surplus of } \{d', n\} \text{ had } d' \text{ employed } n \text{ in } t \text{ (counterfactual)}} \\ & \underbrace{- \delta[1-\eta(d, \kappa_{n,t})] \int_{\epsilon_{n,t+1}(e)} \mathbb{E} V_{d'}(s_{n,t+1}(e), \epsilon_{n,t+1}(e) | s_{n,t}(e); d) dG_e}_{\text{EPDV of match surplus of } \{d', n\} \text{ when } d \text{ employs } n \text{ in } t}\end{aligned}$$

- This wage equation follows immediately from worker indifference and non-employing firm indifference

# Econometric Framework

# Dynamic Equilibrium Generalized Roy Model

- This framework is a *dynamic generalized equilibrium Roy model*

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{\underbrace{D_{n,t} = d, D'_{n,t} = d'}_{\text{selection}}, e_n = e\} \left[\underbrace{y(d', s_{n,t}(e)) + \Psi(d, d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{potential equilibrium wage } w_{n,t}(d,d',e)}\right]$$

- Given panel of data on wages, employment choices, initial attributes: $(w_{n,t}, D_{n,t}, H_{n,1})$ for $t = 1, \ldots, T$
- Minimal data requirements: no proxies for beliefs or direct information on performance signals
- Even as we allow $D_{n,t}$ and $D'_{n,t}$ to be function of *all* variables affecting potential wages
- Including unobservables $P_{n,t}, e_n$ and $\epsilon_{n,t}$
  - ▷ So *dynamic selection on unobservables* naturally arises
  - ▷ Based on time-varying $(P_{n,t}, \epsilon_{n,t})$ and endogenously evolving $P_{n,t}$

# Dynamic Equilibrium Generalized Roy Model

- This framework is a *dynamic generalized equilibrium Roy model*

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{\underbrace{D_{n,t} = d, D'_{n,t} = d'}_{\text{selection}}, e_n = e\} \left[\underbrace{y(d', s_{n,t}(e)) + \Psi(d, d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{potential equilibrium wage } w_{n,t}(d,d',e)}\right]$$

- Given panel of data on wages, employment choices, initial attributes: $(w_{n,t}, D_{n,t}, H_{n,1})$ for $t = 1, \ldots, T$
- Minimal data requirements: no proxies for beliefs or direct information on performance signals
- Even as we allow $D_{n,t}$ and $D'_{n,t}$ to be function of *all* variables affecting potential wages
- Including unobservables $P_{n,t}, e_n$ and $\epsilon_{n,t}$
  - ▷ So *dynamic selection on unobservables* naturally arises
  - ▷ Based on time-varying $(P_{n,t}, \epsilon_{n,t})$ and endogenously evolving $P_{n,t}$

# Dynamic Equilibrium Generalized Roy Model

- This framework is a *dynamic generalized equilibrium Roy model*

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{\underbrace{D_{n,t} = d, D'_{n,t} = d'}_{\text{selection}}, e_n = e\} \left[\underbrace{y(d', s_{n,t}(e)) + \Psi(d, d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{potential equilibrium wage } w_{n,t}(d,d',e)}\right]$$

- Given panel of data on wages, employment choices, initial attributes: $(w_{n,t}, D_{n,t}, H_{n,1})$ for $t = 1, \ldots, T$
- Minimal data requirements: no proxies for beliefs or direct information on performance signals
- Even as we allow $D_{n,t}$ and $D'_{n,t}$ to be function of *all* variables affecting potential wages
- Including unobservables $P_{n,t}, e_n$ and $\epsilon_{n,t}$
  - ▷ So *dynamic selection on unobservables* naturally arises
  - ▷ Based on time-varying $(P_{n,t}, \epsilon_{n,t})$ and endogenously evolving $P_{n,t}$

# Dynamic Equilibrium Generalized Roy Model

- This framework is a *dynamic generalized equilibrium Roy model*

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{\underbrace{D_{n,t} = d, D'_{n,t} = d'}_{\text{selection}}, e_n = e\} \left[\underbrace{y(d', s_{n,t}(e)) + \Psi(d, d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{potential equilibrium wage } w_{n,t}(d,d',e)}\right]$$

- Given panel of data on wages, employment choices, initial attributes: $(w_{n,t}, D_{n,t}, H_{n,1})$ for $t = 1, \ldots, T$
- Minimal data requirements: no proxies for beliefs or direct information on performance signals
- Even as we allow $D_{n,t}$ and $D'_{n,t}$ to be function of *all* variables affecting potential wages
- Including unobservables $P_{n,t}, e_n$ and $\epsilon_{n,t}$
  - ▷ So *dynamic selection on unobservables* naturally arises
  - ▷ Based on time-varying $(P_{n,t}, \epsilon_{n,t})$ and endogenously evolving $P_{n,t}$

# Dynamic Equilibrium Generalized Roy Model

- This framework is a *dynamic generalized equilibrium Roy model*

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{\underbrace{D_{n,t} = d, D'_{n,t} = d'}_{\text{selection}}, e_n = e\} \Big[\underbrace{y(d', s_{n,t}(e)) + \Psi(d, d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{potential equilibrium wage } w_{n,t}(d,d',e)}\Big]$$

- Given panel of data on wages, employment choices, initial attributes: $(w_{n,t}, D_{n,t}, H_{n,1})$ for $t = 1, \ldots, T$
- Minimal data requirements: no proxies for beliefs or direct information on performance signals
- Even as we allow $D_{n,t}$ and $D'_{n,t}$ to be function of *all* variables affecting potential wages
- Including unobservables $P_{n,t}, e_n$ and $\epsilon_{n,t}$
  - ▷ So *dynamic selection on unobservables* naturally arises
  - ▷ Based on time-varying $(P_{n,t}, \epsilon_{n,t})$ and endogenously evolving $P_{n,t}$

# Dynamic Equilibrium Generalized Roy Model

- This framework is a *dynamic generalized equilibrium Roy model*

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{\underbrace{D_{n,t} = d, D'_{n,t} = d'}_{\text{selection}}, e_n = e\} \Big[\underbrace{y(d', s_{n,t}(e)) + \Psi(d, d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{potential equilibrium wage } w_{n,t}(d,d',e)}\Big]$$

- Given panel of data on wages, employment choices, initial attributes: $(w_{n,t}, D_{n,t}, H_{n,1})$ for $t = 1, \ldots, T$
- Minimal data requirements: no proxies for beliefs or direct information on performance signals
- Even as we allow $D_{n,t}$ and $D'_{n,t}$ to be function of *all* variables affecting potential wages
- Including unobservables $P_{n,t}, e_n$ and $\epsilon_{n,t}$
  - ▷ So *dynamic selection on unobservables* naturally arises
  - ▷ Based on time-varying $(P_{n,t}, \epsilon_{n,t})$ and endogenously evolving $P_{n,t}$

# Dynamic Equilibrium Generalized Roy Model

- This framework is a *dynamic generalized equilibrium Roy model*

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{\underbrace{D_{n,t} = d, D'_{n,t} = d'}_{\text{selection}}, e_n = e\} \Big[\underbrace{y(d', s_{n,t}(e)) + \Psi(d, d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{potential equilibrium wage } w_{n,t}(d,d',e)}\Big]$$

- Given panel of data on wages, employment choices, initial attributes: $(w_{n,t}, D_{n,t}, H_{n,1})$ for $t = 1, \ldots, T$
- Minimal data requirements: no proxies for beliefs or direct information on performance signals
- Even as we allow $D_{n,t}$ and $D'_{n,t}$ to be function of *all* variables affecting potential wages
- Including unobservables $P_{n,t}, e_n$ and $\epsilon_{n,t}$
  - ▷ So *dynamic selection on unobservables* naturally arises
  - ▷ Based on time-varying $(P_{n,t}, \epsilon_{n,t})$ and endogenously evolving $P_{n,t}$

# Dynamic Equilibrium Generalized Roy Model

- This framework is a *dynamic generalized equilibrium Roy model*

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{\underbrace{D_{n,t} = d, D'_{n,t} = d'}_{\text{selection}}, e_n = e\} \Big[\underbrace{y(d', s_{n,t}(e)) + \Psi(d, d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)}_{\text{potential equilibrium wage } w_{n,t}(d,d',e)}\Big]$$

- Given panel of data on wages, employment choices, initial attributes: $(w_{n,t}, D_{n,t}, H_{n,1})$ for $t = 1, \ldots, T$
- Minimal data requirements: no proxies for beliefs or direct information on performance signals
- Even as we allow $D_{n,t}$ and $D'_{n,t}$ to be function of *all* variables affecting potential wages
- Including unobservables $P_{n,t}, e_n$ and $\epsilon_{n,t}$
  - ▷ So *dynamic selection on unobservables* naturally arises
  - ▷ Based on time-varying $(P_{n,t}, \epsilon_{n,t})$ and endogenously evolving $P_{n,t}$

# Primitives to Identify

- To study the impact of firm-worker sorting on earnings inequality, we identify
  - ▷ "Deterministic" wage component $\varphi(\cdot) := y(\cdot) + \Psi(\cdot)$
  - ▷ Output/HK technology as captured by $y(\cdot)$ and compensating differential $\Psi(\cdot)$
  - ▷ Information technology described by prior and signal distribution
  - ▷ Distribution of $\epsilon_{n,t}$

- We identify other important primitives of dynamic models namely
  - ▷ Law of motion of the state $s_{n,t}(e)$
  - ▷ Conditional choice probabilities (CCPs)

# Primitives to Identify

- To study the impact of firm-worker sorting on earnings inequality, we identify
  - ▷ "Deterministic" wage component $\varphi(\cdot) := y(\cdot) + \Psi(\cdot)$
  - ▷ Output/HK technology as captured by $y(\cdot)$ and compensating differential $\Psi(\cdot)$
  - ▷ Information technology described by prior and signal distribution
  - ▷ Distribution of $\epsilon_{n,t}$

- We identify other important primitives of dynamic models namely
  - ▷ Law of motion of the state $s_{n,t}(e)$
  - ▷ Conditional choice probabilities (CCPs)

# Identification

# Challenges to Identification

- Dynamic generalized equilibrium Roy model

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{D_{n,t} = d, D'_{n,t} = d', e_n = e\} \big[ \underbrace{\varphi(d, d', s_{n,t}(e))}_{y(\cdot)+\Psi(\cdot)} + \epsilon_{n,t}(d', e) \big]$$

- Standard challenge: with $D'_{n,t}$ and $s_{n,t}(e)$ observed, identifying $\varphi(\cdot)$ contaminated by selection on $\epsilon_{n,t}$

$$\mathbb{E}[w_{n,t} \mid D_{n,t} = d, D'_{n,t} = d', s_{n,t}(e)] = \varphi(d, d', s_{n,t}(e)) + \underbrace{\mathbb{E}(\epsilon_{n,t}(d', e) \mid D_{n,t} = d, D'_{n,t} = d', s_{n,t}(e))}_{\lambda(d, d', s_{n,t}(e)):\ \text{selection bias}}$$

- Additional challenge: $D'_{n,t}$ and $s_{n,t}(e)$ are unobserved so $\mathbb{E}[w_{n,t} \mid D_{n,t} = d, D'_{n,t} = d', s_{n,t}(e)]$ is unknown

- For now, set aside challenges due to the non-observability of 2nd-best firm $D'_{n,t}$
  - ▷ Assume $D'_{n,t}$ is known
  - ▷ Identification argument immediately extends to the case when $D'_{n,t}$ is unknown

# Challenges to Identification

- Dynamic generalized equilibrium Roy model

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{D_{n,t} = d, D'_{n,t} = d', e_n = e\} \left[ \underbrace{\varphi(d, d', s_{n,t}(e))}_{y(\cdot)+\Psi(\cdot)} + \epsilon_{n,t}(d', e) \right]$$

- Standard challenge: with $D'_{n,t}$ and $s_{n,t}(e)$ observed, identifying $\varphi(\cdot)$ contaminated by selection on $\epsilon_{n,t}$

$$\mathbb{E}[w_{n,t} \mid D_{n,t} = d,\, D'_{n,t} = d', s_{n,t}(e)] = \varphi(d, d', s_{n,t}(e)) + \underbrace{\mathbb{E}(\epsilon_{n,t}(d', e) \mid D_{n,t} = d,\, D'_{n,t} = d', s_{n,t}(e))}_{\lambda(d, d', s_{n,t}(e)):\ \text{selection bias}}$$

- Additional challenge: $D'_{n,t}$ and $s_{n,t}(e)$ are unobserved so $\mathbb{E}[w_{n,t} \mid D_{n,t} = d,\, D'_{n,t} = d',\, s_{n,t}(e)]$ is unknown

- For now, set aside challenges due to the non-observability of 2nd-best firm $D'_{n,t}$
  - ▷ Assume $D'_{n,t}$ is known
  - ▷ Identification argument immediately extends to the case when $D'_{n,t}$ is unknown [Roy]

# Challenges to Identification

- Dynamic generalized equilibrium Roy model

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{D_{n,t} = d, D'_{n,t} = d', e_n = e\}\left[\underbrace{\varphi(d, d', s_{n,t}(e))}_{y(\cdot)+\Psi(\cdot)} + \epsilon_{n,t}(d', e)\right]$$

- Standard challenge: with $D'_{n,t}$ and $s_{n,t}(e)$ observed, identifying $\varphi(\cdot)$ contaminated by selection on $\epsilon_{n,t}$

$$\mathbb{E}[w_{n,t} \mid D_{n,t} = d,\, D'_{n,t} = d', s_{n,t}(e)] = \varphi(d, d', s_{n,t}(e)) + \underbrace{\mathbb{E}(\epsilon_{n,t}(d', e) \mid D_{n,t} = d,\, D'_{n,t} = d', s_{n,t}(e))}_{\lambda(d,d',s_{n,t}(e)):\ \text{selection bias}}$$

- Additional challenge: $D'_{n,t}$ and $s_{n,t}(e)$ are unobserved so $\mathbb{E}[w_{n,t} \mid D_{n,t} = d,\, D'_{n,t} = d', s_{n,t}(e)]$ is unknown

- For now, set aside challenges due to the non-observability of 2nd-best firm $D'_{n,t}$
  - ▷ Assume $D'_{n,t}$ is known
  - ▷ Identification argument immediately extends to the case when $D'_{n,t}$ is unknown [Roy]

# Challenges to Identification

- Dynamic generalized equilibrium Roy model

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{D_{n,t} = d, D'_{n,t} = d', e_n = e\}\left[\underbrace{\varphi(d, d', s_{n,t}(e))}_{y(\cdot)+\Psi(\cdot)} + \epsilon_{n,t}(d', e)\right]$$

- Standard challenge: with $D'_{n,t}$ and $s_{n,t}(e)$ observed, identifying $\varphi(\cdot)$ contaminated by selection on $\epsilon_{n,t}$

$$\mathbb{E}[w_{n,t} \mid D_{n,t} = d,\, D'_{n,t} = d', s_{n,t}(e)] = \varphi(d, d', s_{n,t}(e)) + \underbrace{\mathbb{E}(\epsilon_{n,t}(d', e) \mid D_{n,t} = d,\, D'_{n,t} = d', s_{n,t}(e))}_{\lambda(d, d', s_{n,t}(e)):\ \text{selection bias}}$$

- Additional challenge: $D'_{n,t}$ and $s_{n,t}(e)$ are unobserved so $\mathbb{E}[w_{n,t} \mid D_{n,t} = d,\, D'_{n,t} = d', s_{n,t}(e)]$ is unknown
- For now, set aside challenges due to the non-observability of 2nd-best firm $D'_{n,t}$
  - ▷ Assume $D'_{n,t}$ is known
  - ▷ Identification argument immediately extends to the case when $D'_{n,t}$ is unknown [Roy]

# Challenges to Identification

- Dynamic generalized equilibrium Roy model

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{D_{n,t} = d, D'_{n,t} = d', e_n = e\} \big[ \underbrace{\varphi(d, d', s_{n,t}(e))}_{y(\cdot)+\Psi(\cdot)} + \epsilon_{n,t}(d', e) \big]$$

- Standard challenge: with $D'_{n,t}$ and $s_{n,t}(e)$ observed, identifying $\varphi(\cdot)$ contaminated by selection on $\epsilon_{n,t}$

$$\mathbb{E}[w_{n,t} \mid D_{n,t} = d, D'_{n,t} = d', s_{n,t}(e)] = \varphi(d, d', s_{n,t}(e)) + \underbrace{\mathbb{E}(\epsilon_{n,t}(d', e) \mid D_{n,t} = d, D'_{n,t} = d', s_{n,t}(e))}_{\lambda(d, d', s_{n,t}(e)):\ \text{selection bias}}$$

- Additional challenge: $D'_{n,t}$ and $s_{n,t}(e)$ are unobserved so $\mathbb{E}[w_{n,t} \mid D_{n,t} = d, D'_{n,t} = d', s_{n,t}(e)]$ is unknown

- For now, set aside challenges due to the non-observability of 2nd-best firm $D'_{n,t}$
  - ▷ Assume $D'_{n,t}$ is known
  - ▷ Identification argument immediately extends to the case when $D'_{n,t}$ is unknown ▸ [Roy]

# Challenges to Identification

- Dynamic generalized equilibrium Roy model

$$w_{n,t} = \sum_{d,d'} \mathbb{1}\{D_{n,t} = d, D'_{n,t} = d', e_n = e\} \Big[ \underbrace{\varphi(d, d', s_{n,t}(e))}_{y(\cdot)+\Psi(\cdot)} + \epsilon_{n,t}(d', e) \Big]$$

- Standard challenge: with $D'_{n,t}$ and $s_{n,t}(e)$ observed, identifying $\varphi(\cdot)$ contaminated by selection on $\epsilon_{n,t}$

$$\mathbb{E}[w_{n,t} \mid D_{n,t} = d,\, D'_{n,t} = d', s_{n,t}(e)] = \varphi(d, d', s_{n,t}(e)) + \underbrace{\mathbb{E}(\epsilon_{n,t}(d', e) \mid D_{n,t} = d,\, D'_{n,t} = d', s_{n,t}(e))}_{\lambda(d, d', s_{n,t}(e)):\ \text{selection bias}}$$

- Additional challenge: $D'_{n,t}$ and $s_{n,t}(e)$ are unobserved so $\mathbb{E}[w_{n,t} \mid D_{n,t} = d,\, D'_{n,t} = d',\, s_{n,t}(e)]$ is unknown

- For now, set aside challenges due to the non-observability of 2nd-best firm $D'_{n,t}$
  - ▷ Assume $D'_{n,t}$ is known
  - ▷ Identification argument immediately extends to the case when $D'_{n,t}$ is unknown [Roy]

# Intuition for Our Approach and Its Novelty

- We show identification through *mixture-based approach* building on arguments for static Roy models

- We *solely* rely on information on job choices and wages

- Our identification arguments *do not* require restrictions on
  - ▷ Endogenous variables (e.g. monotonicity restrictions)
  - ▷ The dynamics of states, choices or outcomes (e.g. "sufficient" job mobility as in AKM)

- Rather, our identification arguments rely on conditions that
  - ▷ Impose minimal data requirements
  - ▷ Allow for arbitrary patterns of selection based on endogenously time-varying unobservables
  - ▷ Are easy to verify
  - ▷ Lead to constructive estimators of primitives

# Intuition for Our Approach and Its Novelty

- We show identification through *mixture-based approach* building on arguments for static Roy models
- We *solely* rely on information on job choices and wages
- Our identification arguments *do not* require restrictions on
  - ▷ Endogenous variables (e.g. monotonicity restrictions)
  - ▷ The dynamics of states, choices or outcomes (e.g. "sufficient" job mobility as in AKM)
- Rather, our identification arguments rely on conditions that
  - ▷ Impose minimal data requirements
  - ▷ Allow for arbitrary patterns of selection based on endogenously time-varying unobservables
  - ▷ Are easy to verify
  - ▷ Lead to constructive estimators of primitives

# Intuition for Our Approach and Its Novelty

- We show identification through *mixture-based approach* building on arguments for static Roy models
- We *solely* rely on information on job choices and wages
- Our identification arguments *do not* require restrictions on
  - ▷ Endogenous variables (e.g. monotonicity restrictions)
  - ▷ The dynamics of states, choices or outcomes (e.g. "sufficient" job mobility as in AKM)
- Rather, our identification arguments rely on conditions that
  - ▷ Impose minimal data requirements
  - ▷ Allow for arbitrary patterns of selection based on endogenously time-varying unobservables
  - ▷ Are easy to verify
  - ▷ Lead to constructive estimators of primitives

# An Overview of Our Identification Approach

1. Identify distribution of wages as nonparametric *wage mixture* (Aragam & al. 2020) by
   - ▷ Representing observed distribution of paid wages $w_{n,t}$ for each job history $D_n^t$ as a mixture over $(e_n, a_n^{t-1})$
   - ▷ Showing identification of mixture weights and components under nonparametric restrictions

2. Concatenate *mixture weights* across periods to identify the distribution of $(P_{n,t}, e_n)$ and $s_{n,t}(P_{n,t+1}, e_n, H_{n,1}, \kappa_{n,t+1})$
   - ▷ How? Recall $P_{n,t}$ is essentially a function of $(e_n, D_n^{t-1}, a_n^{t-1})$
   - ▷ Hence, by combining mixture weights across periods, we identify the distribution of $P_{n,t}$

3. Combine 2. and *mixture components* to identify the distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$

4. Adapt extremal quantile regression arguments to identify deterministic wage $\varphi(\cdot) := y(\cdot) + \Psi(\cdot)$
   - ▷ Chernozhukov, 2005; D'Haultfœuille-Maurel, 2013

5. Combine $\varphi(\cdot)$ with knowledge of wage mixture to identify distribution of $\epsilon_{n,t}$

6. Once *mixture weights* are concatenated across periods, also identify law of motion of the state and CCPs

7. Exploit dynamic *discrete* choice logic to identify $y(\cdot)$ and, in turn, separate $\Psi(\cdot)$ from $\varphi(\cdot)$ (e.g. Magnac-Thesmar, 2002)

# An Overview of Our Identification Approach

1. Identify distribution of wages as nonparametric *wage mixture* (Aragam & al. 2020) by
   - ▷ Representing observed distribution of paid wages $w_{n,t}$ for each job history $D_n^t$ as a mixture over $(e_n, a_n^{t-1})$
   - ▷ Showing identification of mixture weights and components under nonparametric restrictions
2. Concatenate *mixture weights* across periods to identify the distribution of $(P_{n,t}, e_n)$ and $s_{n,t}(P_{n,t+1}, e_n, H_{n,1}, \kappa_{n,t+1})$
   - ▷ How? Recall $P_{n,t}$ is essentially a function of $(e_n, D_n^{t-1}, a_n^{t-1})$
   - ▷ Hence, by combining mixture weights across periods, we identify the distribution of $P_{n,t}$
3. Combine 2. and *mixture components* to identify the distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$
4. Adapt extremal quantile regression arguments to identify deterministic wage $\varphi(\cdot) := y(\cdot) + \Psi(\cdot)$
   - ▷ Chernozhukov, 2005; D'Haultfœuille-Maurel, 2013
5. Combine $\varphi(\cdot)$ with knowledge of wage mixture to identify distribution of $\epsilon_{n,t}$
6. Once *mixture weights* are concatenated across periods, also identify law of motion of the state and CCPs
7. Exploit dynamic *discrete* choice logic to identify $y(\cdot)$ and, in turn, separate $\Psi(\cdot)$ from $\varphi(\cdot)$ (e.g. Magnac-Thesmar, 2002)

# An Overview of Our Identification Approach

1. Identify distribution of wages as nonparametric *wage mixture* (Aragam & al. 2020) by
   - ▷ Representing observed distribution of paid wages $w_{n,t}$ for each job history $D_n^t$ as a mixture over $(e_n, a_n^{t-1})$
   - ▷ Showing identification of mixture weights and components under nonparametric restrictions
2. Concatenate *mixture weights* across periods to identify the distribution of $(P_{n,t}, e_n)$ and $s_{n,t}(P_{n,t+1}, e_n, H_{n,1}, \kappa_{n,t+1})$
   - ▷ How? Recall $P_{n,t}$ is essentially a function of $(e_n, D_n^{t-1}, a_n^{t-1})$
   - ▷ Hence, by combining mixture weights across periods, we identify the distribution of $P_{n,t}$
3. Combine 2. and *mixture components* to identify the distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$
4. Adapt extremal quantile regression arguments to identify deterministic wage $\varphi(\cdot) := y(\cdot) + \Psi(\cdot)$
   - ▷ Chernozhukov, 2005; D'Haultfœuille-Maurel, 2013
5. Combine $\varphi(\cdot)$ with knowledge of wage mixture to identify distribution of $\epsilon_{n,t}$
6. Once *mixture weights* are concatenated across periods, also identify law of motion of the state and CCPs
7. Exploit dynamic *discrete* choice logic to identify $y(\cdot)$ and, in turn, separate $\Psi(\cdot)$ from $\varphi(\cdot)$ (e.g. Magnac-Thesmar, 2002)

# Empirical Exercise

# Question

- Reevaluate impact of sorting of high-wage workers into high-paying firms on U.S. earnings inequality
- Influential framework used to analyze earnings inequality in several countries is AKM
  - ▷ Impact of sorting on inequality estimated as *fraction of variance explained by firm-worker effects covariance*
- These estimates often imply *small* impact of sorting as suggested by weak firm-worker effects correlations
  - ▷ Except for Bonhomme & al. (2023): bias correction methods increase importance of sorting
- In our framework: compensating differential and endogenous matching frictions reduce measured sorting
- Two approaches: simulation-based and estimation with LEHD (U.S. matched employer-employee) data

# Question

- Reevaluate impact of sorting of high-wage workers into high-paying firms on U.S. earnings inequality
- Influential framework used to analyze earnings inequality in several countries is AKM
  - ▷ Impact of sorting on inequality estimated as *fraction of variance explained by firm-worker effects covariance*
- These estimates often imply *small* impact of sorting as suggested by weak firm-worker effects correlations
  - ▷ Except for Bonhomme & al. (2023): bias correction methods increase importance of sorting
- In our framework: compensating differential and endogenous matching frictions reduce measured sorting
- Two approaches: simulation-based and estimation with LEHD (U.S. matched employer-employee) data

# Question

- Reevaluate impact of sorting of high-wage workers into high-paying firms on U.S. earnings inequality
- Influential framework used to analyze earnings inequality in several countries is AKM
  - ▷ Impact of sorting on inequality estimated as *fraction of variance explained by firm-worker effects covariance*
- These estimates often imply *small* impact of sorting as suggested by weak firm-worker effects correlations
  - ▷ Except for Bonhomme & al. (2023): bias correction methods increase importance of sorting
- In our framework: compensating differential and endogenous matching frictions reduce measured sorting
- Two approaches: simulation-based and estimation with LEHD (U.S. matched employer-employee) data

# Question

- Reevaluate impact of sorting of high-wage workers into high-paying firms on U.S. earnings inequality
- Influential framework used to analyze earnings inequality in several countries is AKM
  - ▷ Impact of sorting on inequality estimated as *fraction of variance explained by firm-worker effects covariance*
- These estimates often imply *small* impact of sorting as suggested by weak firm-worker effects correlations
  - ▷ Except for Bonhomme & al. (2023): bias correction methods increase importance of sorting
- In our framework: compensating differential and endogenous matching frictions reduce measured sorting
- Two approaches: simulation-based and estimation with LEHD (U.S. matched employer-employee) data

# Question

- Reevaluate impact of sorting of high-wage workers into high-paying firms on U.S. earnings inequality
- Influential framework used to analyze earnings inequality in several countries is AKM
  - ▷ Impact of sorting on inequality estimated as *fraction of variance explained by firm-worker effects covariance*
- These estimates often imply *small* impact of sorting as suggested by weak firm-worker effects correlations
  - ▷ Except for Bonhomme & al. (2023): bias correction methods increase importance of sorting
- In our framework: compensating differential and endogenous matching frictions reduce measured sorting
- Two approaches: simulation-based and estimation with LEHD (U.S. matched employer-employee) data

# Question

- Reevaluate impact of sorting of high-wage workers into high-paying firms on U.S. earnings inequality
- Influential framework used to analyze earnings inequality in several countries is AKM
  - ▷ Impact of sorting on inequality estimated as *fraction of variance explained by firm-worker effects covariance*
- These estimates often imply *small* impact of sorting as suggested by weak firm-worker effects correlations
  - ▷ Except for Bonhomme & al. (2023): bias correction methods increase importance of sorting
- In our framework: compensating differential and endogenous matching frictions reduce measured sorting
- Two approaches: simulation-based and estimation with LEHD (U.S. matched employer-employee) data

# Preview of Results: Simulation-Based Evidence

- Simulate simplified version of our model that replicates key features of U.S. data
  - ▷ Its parameters set to match U.S. earnings moments (PSID) and AKM-type moments (Song & al. 2019)

- When $\Psi(\cdot)$ is *negative*/*positive*, AKM variance decomposition *dampens*/*amplify* estimated sorting
  - ▷ Relative to the case when $\Psi(\cdot) = 0$

- Because AKM omits compensating differential $\Psi(\cdot)$ and conflates it with $\epsilon_{n,t}$
  - ▷ Leading to a form of "omitted variable bias"

# Preview of Results: Empirical Evidence

- Estimate our wage equation using U.S. Census data (LEHD)

- Finding: we estimate $\Psi(\cdot)$ to be *negative*

- Implies that AKM decomposition *under*-estimates impact of sorting because it omits $\Psi(\cdot)$

- Corroborate these findings with alternative measure of impact of sorting based on *random worker reallocation*

# Simulation: Data Construction

- Simulate an economy replicating key features of U.S. data
- With a few simplifications to facilitate comparison with the AKM framework (still show $\Psi(\cdot)$ key)
  - ▷ Suppress dependence of wages on 2nd-best firm
  - ▷ Assume away Roy selection on shocks $\epsilon_{n,t}$
  - ▷ As in Bonhomme & al. (2019), consider finite number of worker and firm types
- Workers earn our (parameterized) equilibrium wage

$$w_{n,t} = \sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\}$$
$$\times \left[e + \beta_0(d) + \beta_1(d,e)H_{n,1} + \beta_2(d,e)\kappa_{n,t} + \beta_3(d,e)P_{n,t} + \Psi(d,e,H_{n,1},\kappa_{n,t},P_{n,t}) + \epsilon_{n,t}(d,e)\right]$$

  - ▷ $y(\cdot)$ containing $e + \beta_0(d)$ (á la AKM) and first-order terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$
  - ▷ $\Psi(\cdot)$ approx. by truncated Taylor series (higher powers and interaction terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$)

# Simulation: Data Construction

- Simulate an economy replicating key features of U.S. data
- With a few simplifications to facilitate comparison with the AKM framework (still show $\Psi(\cdot)$ key)
  - ▷ Suppress dependence of wages on 2nd-best firm
  - ▷ Assume away Roy selection on shocks $\epsilon_{n,t}$
  - ▷ As in Bonhomme & al. (2019), consider finite number of worker and firm types
- Workers earn our (parameterized) equilibrium wage

$$w_{n,t} = \sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\} \\ \times \left[ e + \beta_0(d) + \beta_1(d,e) H_{n,1} + \beta_2(d,e)\kappa_{n,t} + \beta_3(d,e) P_{n,t} + \Psi(d,e,H_{n,1},\kappa_{n,t},P_{n,t}) + \epsilon_{n,t}(d,e) \right]$$

  - ▷ $y(\cdot)$ containing $e + \beta_0(d)$ (á la AKM) and first-order terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$
  - ▷ $\Psi(\cdot)$ approx. by truncated Taylor series (higher powers and interaction terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$)

# Simulation: Data Construction

- Simulate an economy replicating key features of U.S. data
- With a few simplifications to facilitate comparison with the AKM framework (still show $\Psi(\cdot)$ key)
  - ▷ Suppress dependence of wages on 2nd-best firm
  - ▷ Assume away Roy selection on shocks $\epsilon_{n,t}$
  - ▷ As in Bonhomme & al. (2019), consider finite number of worker and firm types
- Workers earn our (parameterized) equilibrium wage

$$w_{n,t} = \sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\} \times \left[e + \beta_0(d) + \beta_1(d,e)H_{n,1} + \beta_2(d,e)\kappa_{n,t} + \beta_3(d,e)P_{n,t} + \Psi(d,e,H_{n,1},\kappa_{n,t},P_{n,t}) + \epsilon_{n,t}(d,e)\right]$$

  - ▷ $y(\cdot)$ containing $e + \beta_0(d)$ (á la AKM) and first-order terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$
  - ▷ $\Psi(\cdot)$ approx. by truncated Taylor series (higher powers and interaction terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$)

# Simulation: Data Construction

- Simulate an economy replicating key features of U.S. data
- With a few simplifications to facilitate comparison with the AKM framework (still show $\Psi(\cdot)$ key)
  - ▷ Suppress dependence of wages on 2nd-best firm
  - ▷ Assume away Roy selection on shocks $\epsilon_{n,t}$
  - ▷ As in Bonhomme & al. (2019), consider finite number of worker and firm types
- Workers earn our (parameterized) equilibrium wage

$$
\begin{aligned}
w_{n,t} = & \sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\} \\
& \times \left[ e + \beta_0(d) + \beta_1(d,e) H_{n,1} + \beta_2(d,e)\kappa_{n,t} + \beta_3(d,e) P_{n,t} + \Psi(d,e,H_{n,1},\kappa_{n,t},P_{n,t}) + \epsilon_{n,t}(d,e) \right]
\end{aligned}
$$

  - ▷ $y(\cdot)$ containing $e + \beta_0(d)$ (á la AKM) and first-order terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$
  - ▷ $\Psi(\cdot)$ approx. by truncated Taylor series (higher powers and interaction terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$)

# Simulation: Data Construction

- Simulate an economy replicating key features of U.S. data
- With a few simplifications to facilitate comparison with the AKM framework (still show $\Psi(\cdot)$ key)
  - ▷ Suppress dependence of wages on 2nd-best firm
  - ▷ Assume away Roy selection on shocks $\epsilon_{n,t}$
  - ▷ As in Bonhomme & al. (2019), consider finite number of worker and firm types
- Workers earn our (parameterized) equilibrium wage

$$w_{n,t} = \sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\}$$
$$\times \left[e + \beta_0(d) + \beta_1(d,e)H_{n,1} + \beta_2(d,e)\kappa_{n,t} + \beta_3(d,e)P_{n,t} + \Psi(d,e,H_{n,1},\kappa_{n,t},P_{n,t}) + \epsilon_{n,t}(d,e)\right]$$

  - ▷ $y(\cdot)$ containing $e + \beta_0(d)$ (á la AKM) and first-order terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$
  - ▷ $\Psi(\cdot)$ approx. by truncated Taylor series (higher powers and interaction terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$)

# Simulation: Data Construction

- Simulate an economy replicating key features of U.S. data
- With a few simplifications to facilitate comparison with the AKM framework (still show $\Psi(\cdot)$ key)
  - ▷ Suppress dependence of wages on 2nd-best firm
  - ▷ Assume away Roy selection on shocks $\epsilon_{n,t}$
  - ▷ As in Bonhomme & al. (2019), consider finite number of worker and firm types
- Workers earn our (parameterized) equilibrium wage

$$\begin{aligned} w_{n,t} = & \sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\} \\ & \times \left[ e + \beta_0(d) + \beta_1(d,e)H_{n,1} + \beta_2(d,e)\kappa_{n,t} + \beta_3(d,e)P_{n,t} + \Psi(d,e,H_{n,1},\kappa_{n,t},P_{n,t}) + \epsilon_{n,t}(d,e) \right] \end{aligned}$$

  - ▷ $y(\cdot)$ containing $e + \beta_0(d)$ (á la AKM) and first-order terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$
  - ▷ $\Psi(\cdot)$ approx. by truncated Taylor series (higher powers and interaction terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$)

# Simulation: Data Construction

- Set the wage and simulation parameters to match earnings moments from U.S. PSID
  - ▷ Panel Study of Income Dynamics (PSID): representative survey of U.S. households since 1968
  - ▷ Includes info on wages, employment status and other observables
  - ▷ Consider wage moments: growth, life-cycle first and higher moments, inequality and concentration, etc.

- We also include as targets AKM-type moments from Song & al. (2019) based on U.S. SSA

# Simulation: Data Construction

- Set the wage and simulation parameters to match earnings moments from U.S. PSID
  - ▷ Panel Study of Income Dynamics (PSID): representative survey of U.S. households since 1968
  - ▷ Includes info on wages, employment status and other observables
  - ▷ Consider wage moments: growth, life-cycle first and higher moments, inequality and concentration, etc.

- We also include as targets AKM-type moments from Song & al. (2019) based on U.S. SSA

# Simulation: Results

- As in AKM, we measure impact of sorting on inequality based on firm/worker complementarities in $y(\cdot)$

$$\rho := \mathrm{Cov}(e_n, \beta_0(D_{n,t}))/\mathrm{Var}(w_{n,t})$$

- Estimate it from wage regression with $\Psi(\cdot) = 0$ and coeffs. on covariates independent of $(d, e)$ as in AKM

$$w_{n,t} = \sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\} \times \left[e + \beta_0(d) + \beta_1\cancel{(d,e)}H_{n,1} + \beta_2\cancel{(d,e)}\kappa_{n,t} + \beta_3\cancel{(d,e)}P_{n,t} + \cancel{\Psi(d, H_{n,1}, \kappa_{n,t}, P_{n,t}, e)} + \epsilon_{n,t}(d,e)\right]$$

# Simulation: Results

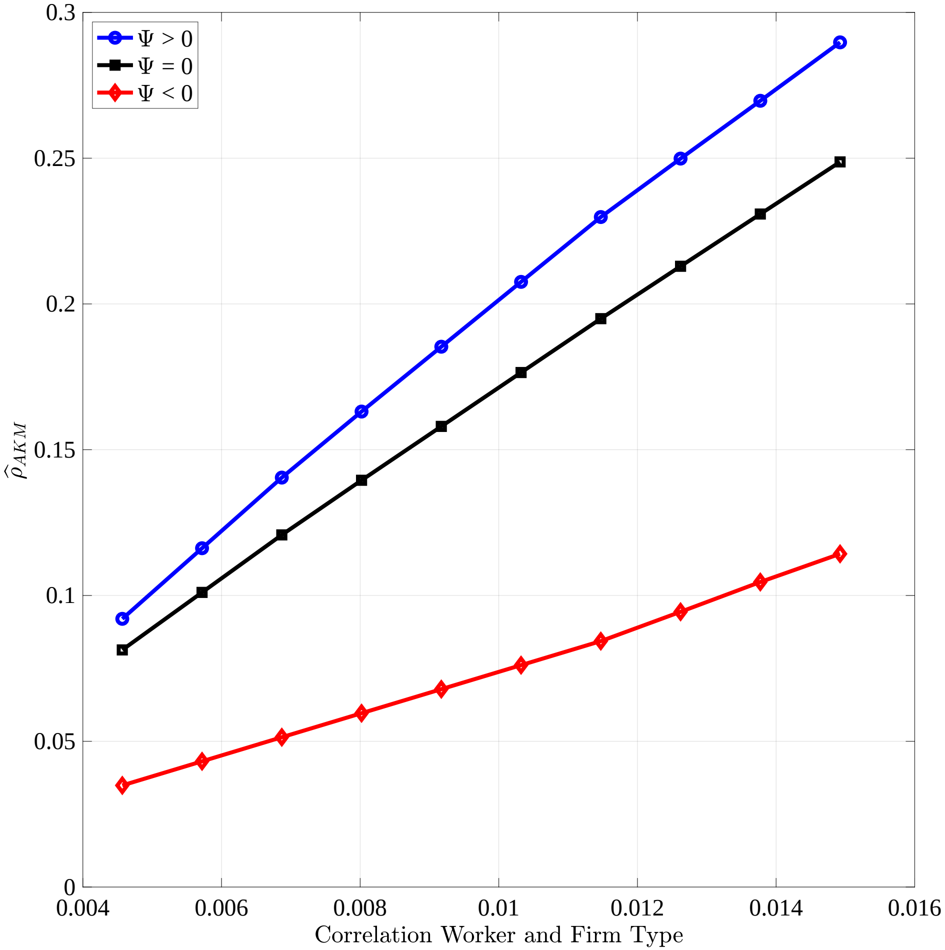
- As in AKM, we measure impact of sorting on inequality based on firm/worker complementarities in $y(\cdot)$

$$\rho := \mathrm{Cov}(e_n, \beta_0(D_{n,t}))/\mathrm{Var}(w_{n,t})$$

- Estimate it from wage regression with $\Psi(\cdot) = 0$ and coeffs. on covariates independent of $(d, e)$ as in AKM

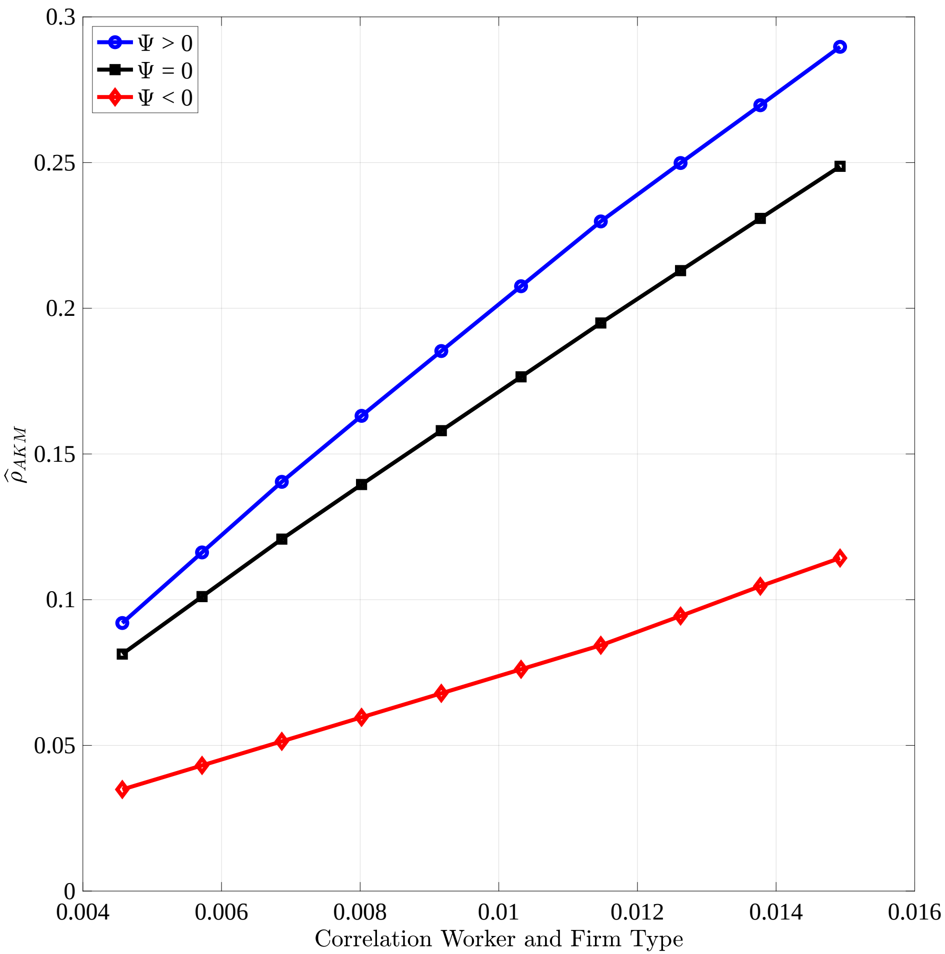
$$w_{n,t} = \sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\}$$
$$\times \left[e + \beta_0(d) + \beta_1\cancel{(d,e)} H_{n,1} + \beta_2\cancel{(d,e)} \kappa_{n,t} + \beta_3\cancel{(d,e)} P_{n,t} + \cancel{\Psi(d, H_{n,1}, \kappa_{n,t}, P_{n,t}, e)} + \epsilon_{n,t}(d,e)\right]$$

# Simulation: Results



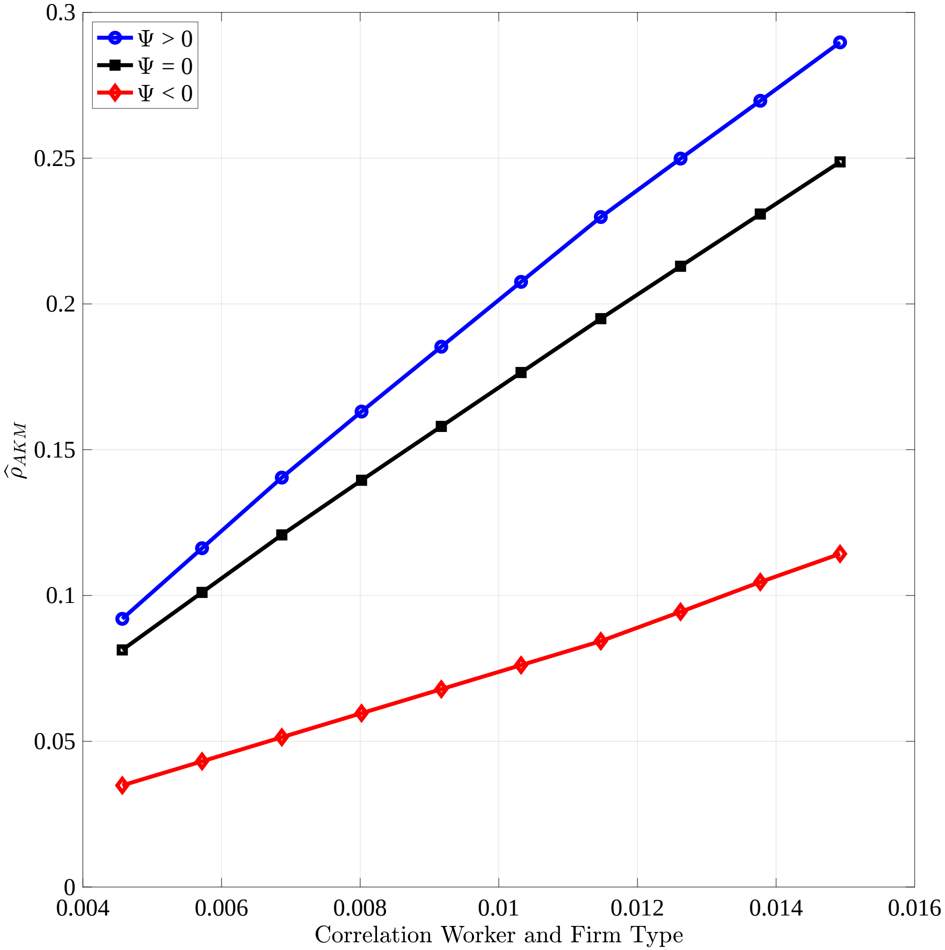

- $\Psi(\cdot) < 0$: workers match with firms offering *higher* HK/info gains than competitors
  - ▷ $\Psi(\cdot) < 0$ *dampens* AKM estimate $\hat{\rho}_{AKM}$ $(< \rho)$
  - ▷ As attenuates firm/worker compl. in output $y(\cdot)$
- $\Psi(\cdot) > 0$: workers match with firms offering *lower* HK/info gains than competitors
  - ▷ $\Psi(\cdot) > 0$ *amplifies* AKM estimate $\hat{\rho}_{AKM}(> \rho)$
  - ▷ As enhances firm/worker compl. in output $y(\cdot)$
- Next: empirical implementation using LEHD data

# Simulation: Results




- $\Psi(\cdot) < 0$: workers match with firms offering *higher* HK/info gains than competitors
  - ▷ $\Psi(\cdot) < 0$ *dampens* AKM estimate $\hat{\rho}_{AKM}$ $(< \rho)$
  - ▷ As attenuates firm/worker compl. in output $y(\cdot)$
- $\Psi(\cdot) > 0$: workers match with firms offering *lower* HK/info gains than competitors
  - ▷ $\Psi(\cdot) > 0$ *amplifies* AKM estimate $\hat{\rho}_{AKM}(> \rho)$
  - ▷ As enhances firm/worker compl. in output $y(\cdot)$
- Next: empirical implementation using LEHD data

# Simulation: Results




- $\Psi(\cdot) < 0$: workers match with firms offering *higher* HK/info gains than competitors
  - ▷ $\Psi(\cdot) < 0$ *dampens* AKM estimate $\hat{\rho}_{AKM}$ $(< \rho)$
  - ▷ As attenuates firm/worker compl. in output $y(\cdot)$

- $\Psi(\cdot) > 0$: workers match with firms offering *lower* HK/info gains than competitors
  - ▷ $\Psi(\cdot) > 0$ *amplifies* AKM estimate $\hat{\rho}_{AKM}(> \rho)$
  - ▷ As enhances firm/worker compl. in output $y(\cdot)$

- Next: empirical implementation using LEHD data

# Empirics: Wage Equation

- Consider equilibrium wage equation parameterized as in the simulations

$$
\begin{aligned}
w_{n,t} = & \sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\} \\
& \times \left[ e + \beta_0(d) + \beta_1(d,e)H_{n,1} + \beta_2(d,e)\kappa_{n,t} + \beta_3(d,e)P_{n,t} + \Psi(d, H_{n,1}, \kappa_{n,t}, P_{n,t}, e) + \epsilon_{n,t}(d,e) \right]
\end{aligned}
$$

- $y(\cdot)$ containing $e + \beta_0(d)$ (á la AKM) and first-order terms of $H_{n,1}, \kappa_{n,t}, P_{n,t}$

- $\Psi(\cdot)$ approx. with truncated Taylor series (higher powers and interaction terms of $H_{n,1}, \kappa_{n,t}$ and $P_{n,t}$)

- Suppress dependence on 2nd-best firm as in AKM but allow for selection on $\epsilon_{n,t}$ as in our class of models

# Empirics: Data

- Estimate wage equation using U.S. LEHD data
  - ▷ LEHD provides administrative data on quarterly labor earnings for all workers across all their jobs
  - ▷ Dataset has info for 21 states (include CA, FL, PA) from 1994 to 2022
  - ▷ Observables: age, gender, education, firm identifier, job location and industry
  - ▷ $w_{n,t}, D_{n,t}, H_{n,1}$ and $\kappa_{n,t}$ are observed in the data

- As for beliefs: simply recover $\{P_{n,t}\}$ process by inferring performance from variable pay in pre-step

- Approach: $P_{n,t}$ is estimated for each $(n, t)$ by extracting performance signals from variable pay (vp)
  - ▷ Why? Do not want to rely on independent measures of performance (often unavailable)
  - ▷ Idea: *quantiles of vp distribution identify performance signals* if vp is monotone with performance
  - ▷ So if worker is in *top quantile of vp distribution*, we infer worker received *high* performance signal

# Empirics: Data

- Estimate wage equation using U.S. LEHD data
  - ▷ LEHD provides administrative data on quarterly labor earnings for all workers across all their jobs
  - ▷ Dataset has info for 21 states (include CA, FL, PA) from 1994 to 2022
  - ▷ Observables: age, gender, education, firm identifier, job location and industry
  - ▷ $w_{n,t}, D_{n,t}, H_{n,1}$ and $\kappa_{n,t}$ are observed in the data
- As for beliefs: simply recover $\{P_{n,t}\}$ process by inferring performance from variable pay in pre-step
- Approach: $P_{n,t}$ is estimated for each $(n,t)$ by extracting performance signals from variable pay (vp)
  - ▷ Why? Do not want to rely on independent measures of performance (often unavailable)
  - ▷ Idea: *quantiles of vp distribution identify performance signals* if vp is monotone with performance
  - ▷ So if worker is in *top quantile of vp distribution*, we infer worker received *high* performance signal

# Empirics: Data

- Estimate wage equation using U.S. LEHD data
  - ▷ LEHD provides administrative data on quarterly labor earnings for all workers across all their jobs
  - ▷ Dataset has info for 21 states (include CA, FL, PA) from 1994 to 2022
  - ▷ Observables: age, gender, education, firm identifier, job location and industry
  - ▷ $w_{n,t}, D_{n,t}, H_{n,1}$ and $\kappa_{n,t}$ are observed in the data

- As for beliefs: simply recover $\{P_{n,t}\}$ process by inferring performance from variable pay in pre-step

- Approach: $P_{n,t}$ is estimated for each $(n, t)$ by extracting performance signals from variable pay (vp)
  - ▷ Why? Do not want to rely on independent measures of performance (often unavailable)
  - ▷ Idea: *quantiles of vp distribution identify performance signals* if vp is monotone with performance
  - ▷ So if worker is in *top quantile of vp distribution*, we infer worker received *high* performance signal

# Empirics: Estimation Method

- Wage equation

$$
\begin{aligned}
w_{n,t} = &\sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\} \\
&\times \left[e + \beta_0(d) + \beta_1(d,e)H_{n,1} + \beta_2(d,e)\kappa_{n,t} + \beta_3(d,e)P_{n,t} + \Psi(d, H_{n,1}, \kappa_{n,t}, P_{n,t}, e) + \epsilon_{n,t}(d,e)\right]
\end{aligned}
$$

- Finite Gaussian mixture in latent worker efficiency types $e_n$
- To address selection on $\epsilon_{n,t}$, build on extremal quantile reg. of D'Haultfœuille-Maurel-Zhang (2018)
- Therefore, nest extremal quantile regression within Gaussian mixture estimation
- Implemented via feasible estimator in the spirit of Stata's `fmm+eqregsel`
- Critical: can estimate key constant term by normalizing error at suitable extremal quantile [Intercept]

# Empirics: Estimation Method

- Wage equation

$$w_{n,t} = \sum_{d,e} \mathbb{1}\{D_{n,t} = d, e_n = e\}$$
$$\times \left[e + \beta_0(d) + \beta_1(d,e)H_{n,1} + \beta_2(d,e)\kappa_{n,t} + \beta_3(d,e)P_{n,t} + \Psi(d, H_{n,1}, \kappa_{n,t}, P_{n,t}, e) + \epsilon_{n,t}(d,e)\right]$$

- Finite Gaussian mixture in latent worker efficiency types $e_n$
- To address selection on $\epsilon_{n,t}$, build on extremal quantile reg. of D'Haultfœuille-Maurel-Zhang (2018)
- Therefore, nest extremal quantile regression within Gaussian mixture estimation
- Implemented via feasible estimator in the spirit of Stata's `fmm`+`eqregsel`
- Critical: can estimate key constant term by normalizing error at suitable extremal quantile [Intercept]

# Empirics: Estimation Results

- AKM decomposition *under*-estimates impact of sorting: $\hat{\rho}_{\text{AKM}} < \hat{\rho}_{\text{OURS}}$
- Because the estimated $\Psi(\cdot)$ is on average *negative*
- As workers tend to match mostly with firms offering HK/info with *high* future wage returns
- Hence, we have gone some way at *solving* the puzzle of low sorting
- Next: corroborate this key finding with exercise capturing *global notion of sorting* in our class of model

# Empirics: Estimation Results

- AKM decomposition *under*-estimates impact of sorting: $\hat{\rho}_{\text{AKM}} < \hat{\rho}_{\text{OURS}}$
- Because the estimated $\Psi(\cdot)$ is on average *negative*
- As workers tend to match mostly with firms offering HK/info with *high* future wage returns
- Hence, we have gone some way at *solving* the puzzle of low sorting
- Next: corroborate this key finding with exercise capturing *global notion of sorting* in our class of model

# Empirics: Estimation Results

- AKM decomposition *under*-estimates impact of sorting: $\hat{\rho}_{\mathsf{AKM}} < \hat{\rho}_{\mathsf{OURS}}$
- Because the estimated $\Psi(\cdot)$ is on average *negative*
- As workers tend to match mostly with firms offering HK/info with *high* future wage returns
- Hence, we have gone some way at *solving* the puzzle of low sorting
- Next: corroborate this key finding with exercise capturing *global notion of sorting* in our class of model

# Role of Sorting: Random-Matching Counterfactual

- Note that $\rho$ (AKM measure of sorting) captures worker sorting *solely* by efficiency $e_n$
- But workers also sort by beliefs about ability $\theta_n$ and accumulated HK (endogenous matching "frictions")
- We capture additional sorting dimensions by comparing earnings moments to random-matching benchmark
- Intuition: *if sorting matters* for inequality then *std. of earnings* must *decrease* under random matching
- Similarly, *if sorting matters* for inequality then *top earnings share* must *decrease* under random matching
- We find evidence of *both* mechanisms
- We confirm that sorting matters for inequality addressing the puzzle that measured sorting is low

# Role of Sorting: Random-Matching Counterfactual

- Note that $\rho$ (AKM measure of sorting) captures worker sorting *solely* by efficiency $e_n$
- But workers also sort by beliefs about ability $\theta_n$ and accumulated HK (endogenous matching "frictions")
- We capture additional sorting dimensions by comparing earnings moments to random-matching benchmark
- Intuition: *if sorting matters* for inequality then *std. of earnings* must *decrease* under random matching
- Similarly, *if sorting matters* for inequality then *top earnings share* must *decrease* under random matching
- We find evidence of *both* mechanisms
- We confirm that sorting matters for inequality addressing the puzzle that measured sorting is low

# Role of Sorting: Random-Matching Counterfactual

- Note that $\rho$ (AKM measure of sorting) captures worker sorting *solely* by efficiency $e_n$
- But workers also sort by beliefs about ability $\theta_n$ and accumulated HK (endogenous matching "frictions")
- We capture additional sorting dimensions by comparing earnings moments to random-matching benchmark
- Intuition: *if sorting matters* for inequality then *std. of earnings* must *decrease* under random matching
- Similarly, *if sorting matters* for inequality then *top earnings share* must *decrease* under random matching
- We find evidence of *both* mechanisms
- We confirm that sorting matters for inequality addressing the puzzle that measured sorting is low

# Conclusion

We provide two sets of results:

1. Develop semiparametric identification arguments for matching models with HK acquisition and learning
   - ▷ Represent wage distribution as a mixture over unobservables
   - ▷ Show identification of wage mixture under mild restrictions
   - ▷ *Recover primitives by combining mixture-based approach w/ quantile methods for generalized Roy*
2. Use it to reevaluate impact of worker sorting into firms on U.S. earnings inequality using LEHD data
   - ▷ Findings: sorting matters but conventional measures of it typically *underestimate* its impact
   - ▷ Because they conflate compensating differential in wages with firm/worker/match effects

# Conclusion

We provide two sets of results:

1. Develop semiparametric identification arguments for matching models with HK acquisition and learning
   - ▷ Represent wage distribution as a mixture over unobservables
   - ▷ Show identification of wage mixture under mild restrictions
   - ▷ *Recover primitives by combining mixture-based approach w/ quantile methods for generalized Roy*

2. Use it to reevaluate impact of worker sorting into firms on U.S. earnings inequality using LEHD data
   - ▷ Findings: sorting matters but conventional measures of it typically *underestimate* its impact
   - ▷ Because they conflate compensating differential in wages with firm/worker/match effects

# Appendix

# Literature on Roy Model: Selection Issue

[Back]

- Simplified static version of our wage equation

$$w_n = \sum_{d\in\{0,1\}} \mathbb{1}\{D_n = d\} w_n(d) \sum_{d\in\{0,1\}} \mathbb{1}\{D_n = d\} \left[ y(d, X_n) + \epsilon_n(d) \right]$$

  - ▷ Two alternatives
  - ▷ No 2nd-best firm, state variables replaced by observed covariates
  - ▷ No subscript $t$, no compensating differential

- *Selection* on $\epsilon_n := (\epsilon_n(1), \epsilon_n(0))$ complicates identification of deterministic wage $y(\cdot)$

$$\mathbb{E}(w_n \mid D_n = d, X_n) = \mathbb{E}(y(d, X_n) + \epsilon_n(d) \mid D_n = d, X_n) = y(d, X_n) + \underbrace{\mathbb{E}(\epsilon_n(d) \mid D_n = d, X_n)}_{\lambda(d, X_n):\ \text{selection bias}}$$

# Literature on Roy Model: Selection Issue [Back]

- Simplified static version of our wage equation

$$w_n = \sum_{d \in \{0,1\}} \mathbb{1}\{D_n = d\} w_n(d) \sum_{d \in \{0,1\}} \mathbb{1}\{D_n = d\} \left[ y(d, X_n) + \epsilon_n(d) \right]$$

  - ▷ Two alternatives
  - ▷ No 2nd-best firm, state variables replaced by observed covariates
  - ▷ No subscript $t$, no compensating differential

- *Selection* on $\epsilon_n := (\epsilon_n(1), \epsilon_n(0))$ complicates identification of deterministic wage $y(\cdot)$

$$\mathbb{E}(w_n \mid D_n = d, X_n) = \mathbb{E}(y(d, X_n) + \epsilon_n(d) \mid D_n = d, X_n) = y(d, X_n) + \underbrace{\mathbb{E}(\epsilon_n(d) \mid D_n = d, X_n)}_{\lambda(d, X_n):\ \text{selection bias}}$$

# Literature on Roy Model: Selection Issue [Back]

- Simplified static version of our wage equation

$$w_n = \sum_{d \in \{0,1\}} \mathbb{1}\{D_n = d\} w_n(d) \sum_{d \in \{0,1\}} \mathbb{1}\{D_n = d\} \left[ y(d, X_n) + \epsilon_n(d) \right]$$

  - ▷ Two alternatives
  - ▷ No 2nd-best firm, state variables replaced by observed covariates
  - ▷ No subscript $t$, no compensating differential

- *Selection* on $\epsilon_n := (\epsilon_n(1), \epsilon_n(0))$ complicates identification of deterministic wage $y(\cdot)$

$$\mathbb{E}(w_n \mid D_n = d, X_n) = \mathbb{E}(y(d, X_n) + \epsilon_n(d) \mid D_n = d, X_n) = y(d, X_n) + \underbrace{\mathbb{E}(\epsilon_n(d) \mid D_n = d, X_n)}_{\lambda(d, X_n):\ \text{selection bias}}$$

# Literature on Roy Model: Selection Issue

[Back]

- Simplified static version of our wage equation

$$w_n = \sum_{d\in\{0,1\}} \mathbb{1}\{D_n = d\} w_n(d) \sum_{d\in\{0,1\}} \mathbb{1}\{D_n = d\} \left[y(d, X_n) + \epsilon_n(d)\right]$$

  - ▷ Two alternatives
  - ▷ No 2nd-best firm, state variables replaced by observed covariates
  - ▷ No subscript $t$, no compensating differential

- *Selection* on $\epsilon_n := (\epsilon_n(1), \epsilon_n(0))$ complicates identification of deterministic wage $y(\cdot)$

$$\mathbb{E}(w_n \mid D_n = d, X_n) = \mathbb{E}(y(d, X_n) + \epsilon_n(d) \mid D_n = d, X_n) = y(d, X_n) + \underbrace{\mathbb{E}(\epsilon_n(d) \mid D_n = d, X_n)}_{\lambda(d, X_n):\ \text{selection bias}}$$

# Literature on Roy Model: Exclusion Restrictions

- Identification with worker-specific excluded covariates affecting job choices and not wages or vice versa
  - ▷ Ahn & Powell (1993), Newey (2009), Das, Newey, & Vella (2003)
- Identification "*at infinity*" with worker-job-specific covariates affecting wage in one job only
  - ▷ Chamberlain (1986), Heckman (1990)
- *Not applicable to our class of models* which lacks *any type* of exclusion restrictions by construction
- All the state variables affect both job choices and wages
- There are no worker-job-specific excluded state variables
  - ▷ Belief? Ability $\theta_n$ is allowed to be general trait and so belief $P_{n,t}$ about $\theta_n$ is not job-specific
  - ▷ Job tenure? HK can depend on entire employment history and so job tenure impacts wages across all jobs

# Literature on Roy Model: Exclusion Restrictions

- Identification with worker-specific excluded covariates affecting job choices and not wages or vice versa
  - ▷ Ahn & Powell (1993), Newey (2009), Das, Newey, & Vella (2003)
- Identification "*at infinity*" with worker-job-specific covariates affecting wage in one job only
  - ▷ Chamberlain (1986), Heckman (1990)
- *Not applicable to our class of models* which lacks *any type* of exclusion restrictions by construction
- All the state variables affect both job choices and wages
- There are no worker-job-specific excluded state variables
  - ▷ Belief? Ability $\theta_n$ is allowed to be general trait and so belief $P_{n,t}$ about $\theta_n$ is not job-specific
  - ▷ Job tenure? HK can depend on entire employment history and so job tenure impacts wages across all jobs

# Literature on Roy Model: Exclusion Restrictions

- Identification with worker-specific excluded covariates affecting job choices and not wages or vice versa
  - ▷ Ahn & Powell (1993), Newey (2009), Das, Newey, & Vella (2003)
- Identification "*at infinity*" with worker-job-specific covariates affecting wage in one job only
  - ▷ Chamberlain (1986), Heckman (1990)
- *Not applicable to our class of models* which lacks *any type* of exclusion restrictions by construction
- All the state variables affect both job choices and wages
- There are no worker-job-specific excluded state variables
  - ▷ Belief? Ability $\theta_n$ is allowed to be general trait and so belief $P_{n,t}$ about $\theta_n$ is not job-specific
  - ▷ Job tenure? HK can depend on entire employment history and so job tenure impacts wages across all jobs

# Literature on Roy Model: Exclusion Restrictions

- Identification with worker-specific excluded covariates affecting job choices and not wages or vice versa
  - ▷ Ahn & Powell (1993), Newey (2009), Das, Newey, & Vella (2003)
- Identification "*at infinity*" with worker-job-specific covariates affecting wage in one job only
  - ▷ Chamberlain (1986), Heckman (1990)
- *Not applicable to our class of models* which lacks *any type* of exclusion restrictions by construction
- All the state variables affect both job choices and wages
- There are no worker-job-specific excluded state variables
  - ▷ Belief? Ability $\theta_n$ is allowed to be general trait and so belief $P_{n,t}$ about $\theta_n$ is not job-specific
  - ▷ Job tenure? HK can depend on entire employment history and so job tenure impacts wages across all jobs

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \; \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1, x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \; \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1, x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \ \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1, x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \; \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1, x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \; \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1, x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \; \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1, x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \; \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1, x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w\to+\infty}\Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \ \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1,x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \; \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1, x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w\to+\infty}\Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0\ \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1,x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \; \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1, x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Literature on Roy Model: No Exclusion Restrictions

- D'Haultfœuille & Maurel (2013): identification "*at infinity*" *without* worker-job-specific excluded covariates
- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then $\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \; \forall x$
  - ▷ Intuition: if selection is truly endogenous, the covariates must be irrelevant in the limit
- It implies no impact of selection on right *extreme tail* of wage distribution and $y(1, x)$ is identified [Details]

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x) \quad \text{as } w \to +\infty$$

- Solves lack of exclusion restrictions but *not* directly applicable to our class of models
  - ▷ Relies on knowledge of distribution of $(D_{n,t}, w_{n,t})$ conditional on $s_{n,t}$ but $s_{n,t}$ partially observed
  - ▷ Requires rethinking about which limit to take given second-price auction-like pricing mechanism [Preview]
  - ▷ Next slides: how we address these two issues and incorporate DM's approach into our framework

# Longitudinal vs. Cross-Sectional Dimension

- Longitudinal dimension to identify learning, state law of motion, CCPs, distribution of $(D_{nt}, w_{nt})$ given $s_{nt}$
  - ▷ By concatenating wage mixture weights across periods
- Based on knowledge of $(D_{nt}, w_{nt})$ given $s_{nt}$, cross-sectional dimension to identify deterministic wage $\varphi(\cdot)$
  - ▷ Apply DM in each period
  - ▷ Why? $\varphi(\cdot)$ is function of $s_{n,t}$ whose support varies across periods due to $\kappa_{n,t}$
  - ▷ $\varphi(\cdot)$ is effectively a time-varying function
  - ▷ Empirics: $\varphi(\cdot)$ is parameterized, longitudinal dimension helps avoid DM's location normalization
- Longitudinal dimension to identify output technology $y(\cdot)$ and compensating differential $\Psi(\cdot)$

# Longitudinal vs. Cross-Sectional Dimension

- Longitudinal dimension to identify learning, state law of motion, CCPs, distribution of $(D_{nt}, w_{nt})$ given $s_{nt}$
  - ▷ By concatenating wage mixture weights across periods
- Based on knowledge of $(D_{nt}, w_{nt})$ given $s_{nt}$, cross-sectional dimension to identify deterministic wage $\varphi(\cdot)$
  - ▷ Apply DM in each period
  - ▷ Why? $\varphi(\cdot)$ is function of $s_{n,t}$ whose support varies across periods due to $\kappa_{n,t}$
  - ▷ $\varphi(\cdot)$ is effectively a time-varying function
  - ▷ Empirics: $\varphi(\cdot)$ is parameterized, longitudinal dimension helps avoid DM's location normalization
- Longitudinal dimension to identify output technology $y(\cdot)$ and compensating differential $\Psi(\cdot)$

# Longitudinal vs. Cross-Sectional Dimension

- Longitudinal dimension to identify learning, state law of motion, CCPs, distribution of $(D_{nt}, w_{nt})$ given $s_{nt}$
  - ▷ By concatenating wage mixture weights across periods
- Based on knowledge of $(D_{nt}, w_{nt})$ given $s_{nt}$, cross-sectional dimension to identify deterministic wage $\varphi(\cdot)$
  - ▷ Apply DM in each period
  - ▷ Why? $\varphi(\cdot)$ is function of $s_{n,t}$ whose support varies across periods due to $\kappa_{n,t}$
  - ▷ $\varphi(\cdot)$ is effectively a time-varying function
  - ▷ Empirics: $\varphi(\cdot)$ is parameterized, longitudinal dimension helps avoid DM's location normalization
- Longitudinal dimension to identify output technology $y(\cdot)$ and compensating differential $\Psi(\cdot)$

# Job Mobility and Job Retention

- Strength of our approach: limited reliance on job mobility
- Some variation in job choices (akin to job mobility) helps identify $y(\cdot)$ and $\Psi(\cdot)$
- Variation in job choices for a given state aids in identifying $y(\cdot)$
- Having workers rank firm $d$ as 1st-best and others as 2nd-best, for the same state, aids in identifying $\Psi(\cdot)$
- Identifying firm's information technology requires observing workers employed at that firm for some periods

# Job Mobility and Job Retention

- Strength of our approach: limited reliance on job mobility
- Some variation in job choices (akin to job mobility) helps identify $y(\cdot)$ and $\Psi(\cdot)$
- Variation in job choices for a given state aids in identifying $y(\cdot)$
- Having workers rank firm $d$ as 1st-best and others as 2nd-best, for the same state, aids in identifying $\Psi(\cdot)$
- Identifying firm's information technology requires observing workers employed at that firm for some periods

# Job Mobility and Job Retention

- Strength of our approach: limited reliance on job mobility
- Some variation in job choices (akin to job mobility) helps identify $y(\cdot)$ and $\Psi(\cdot)$
- Variation in job choices for a given state aids in identifying $y(\cdot)$
- Having workers rank firm $d$ as 1st-best and others as 2nd-best, for the same state, aids in identifying $\Psi(\cdot)$
- Identifying firm's information technology requires observing workers employed at that firm for some periods

# Job Mobility and Job Retention

- Strength of our approach: limited reliance on job mobility
- Some variation in job choices (akin to job mobility) helps identify $y(\cdot)$ and $\Psi(\cdot)$
- Variation in job choices for a given state aids in identifying $y(\cdot)$
- Having workers rank firm $d$ as 1st-best and others as 2nd-best, for the same state, aids in identifying $\Psi(\cdot)$
- Identifying firm's information technology requires observing workers employed at that firm for some periods

# Second-Best Firm

- Assume that *two* firms make wage offers to worker $n$ in each period whose identities depend on $s_{n,t}$ *only*
- Aligns with practical reality: workers typically receive wage offers from a limited number of firms
- Similar to search models which typically assume "incumbent" and "competitor"
- Key econometric implication: when conditioning on $(D_{n,t}, s_{n,t})$ $D'_{n,t}$ is *degenerate* at one point
- Most of our results do not require to know (degenerate) support of $D'_{n,t}$ conditional on $(D_{n,t}, s_{n,t})$
  - ▷ If unknown: identify deterministic wage, output/HK/information technology, CCPs, state law of motion
  - ▷ If known: identify compensating differential

# Second-Best Firm

- Assume that *two* firms make wage offers to worker $n$ in each period whose identities depend on $s_{n,t}$ *only*
- Aligns with practical reality: workers typically receive wage offers from a limited number of firms
- Similar to search models which typically assume "incumbent" and "competitor"
- Key econometric implication: when conditioning on $(D_{n,t}, s_{n,t})$ $D'_{n,t}$ is *degenerate* at one point
- Most of our results do not require to know (degenerate) support of $D'_{n,t}$ conditional on $(D_{n,t}, s_{n,t})$
  - ▷ If unknown: identify deterministic wage, output/HK/information technology, CCPs, state law of motion
  - ▷ If known: identify compensating differential

# Second-Best Firm

- Assume that *two* firms make wage offers to worker $n$ in each period whose identities depend on $s_{n,t}$ *only*
- Aligns with practical reality: workers typically receive wage offers from a limited number of firms
- Similar to search models which typically assume "incumbent" and "competitor"
- Key econometric implication: when conditioning on $(D_{n,t}, s_{n,t})$ $D'_{n,t}$ is *degenerate* at one point
- Most of our results do not require to know (degenerate) support of $D'_{n,t}$ conditional on $(D_{n,t}, s_{n,t})$
  - ▷ If unknown: identify deterministic wage, output/HK/information technology, CCPs, state law of motion
  - ▷ If known: identify compensating differential

# Second-Best Firm

- Assume that *two* firms make wage offers to worker $n$ in each period whose identities depend on $s_{n,t}$ *only*
- Aligns with practical reality: workers typically receive wage offers from a limited number of firms
- Similar to search models which typically assume "incumbent" and "competitor"
- Key econometric implication: when conditioning on $(D_{n,t}, s_{n,t})$ $D'_{n,t}$ is *degenerate* at one point
- Most of our results do not require to know (degenerate) support of $D'_{n,t}$ conditional on $(D_{n,t}, s_{n,t})$
  - ▷ If unknown: identify deterministic wage, output/HK/information technology, CCPs, state law of motion
  - ▷ If known: identify compensating differential

# Second-Best Firm

- Can dispense with the assumption above and allow offers from all firms
- Nonparametrically identify workers' acceptance strategy which in turn pinpoints identity of 2nd-best firm
- Operationalized based on workers' observed transition patterns
- Take worker $n$ employed by $d$ at time $t$ with $D_n^t = d^t$
- Consider group of workers sharing same observed characteristics as worker $n$
- And transiting across the same jobs as $d^t$, following same sequence but potentially different lengths
- Examine jobs prior to reaching $d$ and rank them based on average worker retention
- The highest-ranked pair is 2nd-best job for worker $n$ [Back]

# Second-Best Firm

- Can dispense with the assumption above and allow offers from all firms
- Nonparametrically identify workers' acceptance strategy which in turn pinpoints identity of 2nd-best firm
- Operationalized based on workers' observed transition patterns
- Take worker $n$ employed by $d$ at time $t$ with $D_n^t = d^t$
- Consider group of workers sharing same observed characteristics as worker $n$
- And transiting across the same jobs as $d^t$, following same sequence but potentially different lengths
- Examine jobs prior to reaching $d$ and rank them based on average worker retention
- The highest-ranked pair is 2nd-best job for worker $n$ [Back]

# DM: From Imperfect Competition to Search

- DM applicability goes beyond our class of imperfectly competitive models
- Consider a standard wage equation of search models inspired by output technology of Bagger & al. (2014)

$$w_{n,t}(d) = \omega \gamma_t(d)^{\alpha} H_{n,1}{}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta) U(H_{n,1})$$

  - ▷ $H_{n,1}$ is HK (known and time-invariant for simplicity)
  - ▷ $\omega$ is workers' bargaining weight (known)
  - ▷ $\gamma_t(d)$ is firm/job productivity (unknown time-varying firm effect)
  - ▷ $\alpha$ and $\beta$ are parameters (firm-invariant for simplicity)
  - ▷ $U(H_{n,1}) = z + \delta \mathbb{E}(f(S(\cdot); \omega, \alpha, \beta), f_{\epsilon})$ is value of unemployment with $S(\cdot)$ match surplus function

# DM: From Imperfect Competition to Search

- DM applicability goes beyond our class of imperfectly competitive models

- Consider a standard wage equation of search models inspired by output technology of Bagger & al. (2014)

$$w_{n,t}(d) = \omega \gamma_t(d)^{\alpha} {H_{n,1}}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta) U(H_{n,1})$$

  - ▷ $H_{n,1}$ is HK (known and time-invariant for simplicity)
  - ▷ $\omega$ is workers' bargaining weight (known)
  - ▷ $\gamma_t(d)$ is firm/job productivity (unknown time-varying firm effect)
  - ▷ $\alpha$ and $\beta$ are parameters (firm-invariant for simplicity)
  - ▷ $U(H_{n,1}) = z + \delta \mathbb{E}(f(S(\cdot); \omega, \alpha, \beta), f_\epsilon)$ is value of unemployment with $S(\cdot)$ match surplus function

# DM: From Imperfect Competition to Search

- DM applicability goes beyond our class of imperfectly competitive models
- Consider a standard wage equation of search models inspired by output technology of Bagger & al. (2014)

$$w_{n,t}(d) = \omega \gamma_t(d)^{\alpha} {H_{n,1}}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta)U(H_{n,1})$$

  ▷ $H_{n,1}$ is HK (known and time-invariant for simplicity)

  ▷ $\omega$ is workers' bargaining weight (known)

  ▷ $\gamma_t(d)$ is firm/job productivity (unknown time-varying firm effect)

  ▷ $\alpha$ and $\beta$ are parameters (firm-invariant for simplicity)

  ▷ $U(H_{n,1}) = z + \delta \mathbb{E}(f(S(\cdot); \omega, \alpha, \beta), f_{\epsilon})$ is value of unemployment with $S(\cdot)$ match surplus function

# DM: From Imperfect Competition to Search

- DM applicability goes beyond our class of imperfectly competitive models

- Consider a standard wage equation of search models inspired by output technology of Bagger & al. (2014)

$$w_{n,t}(d) = \omega\gamma_t(d)^{\alpha} H_{n,1}{}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta)U(H_{n,1})$$

  - ▷ $H_{n,1}$ is HK (known and time-invariant for simplicity)
  - ▷ $\omega$ is workers' bargaining weight (known)
  - ▷ $\gamma_t(d)$ is firm/job productivity (unknown time-varying firm effect)
  - ▷ $\alpha$ and $\beta$ are parameters (firm-invariant for simplicity)
  - ▷ $U(H_{n,1}) = z + \delta\mathbb{E}(f(S(\cdot);\omega,\alpha,\beta), f_{\epsilon})$ is value of unemployment with $S(\cdot)$ match surplus function

# DM: From Imperfect Competition to Search

- DM applicability goes beyond our class of imperfectly competitive models
- Consider a standard wage equation of search models inspired by output technology of Bagger & al. (2014)

$$w_{n,t}(d) = \omega \gamma_t(d)^{\alpha} H_{n,1}{}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta) U(H_{n,1})$$

  - ▷ $H_{n,1}$ is HK (known and time-invariant for simplicity)
  - ▷ $\omega$ is workers' bargaining weight (known)
  - ▷ $\gamma_t(d)$ is firm/job productivity (unknown time-varying firm effect)
  - ▷ $\alpha$ and $\beta$ are parameters (firm-invariant for simplicity)
  - ▷ $U(H_{n,1}) = z + \delta \mathbb{E}(f(S(\cdot); \omega, \alpha, \beta), f_{\epsilon})$ is value of unemployment with $S(\cdot)$ match surplus function

# DM: From Imperfect Competition to Search

- DM applicability goes beyond our class of imperfectly competitive models
- Consider a standard wage equation of search models inspired by output technology of Bagger & al. (2014)

$$w_{n,t}(d) = \omega \gamma_t(d)^{\alpha} H_{n,1}{}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta)U(H_{n,1})$$

  - ▷ $H_{n,1}$ is HK (known and time-invariant for simplicity)
  - ▷ $\omega$ is workers' bargaining weight (known)
  - ▷ $\gamma_t(d)$ is firm/job productivity (unknown time-varying firm effect)
  - ▷ $\alpha$ and $\beta$ are parameters (firm-invariant for simplicity)
  - ▷ $U(H_{n,1}) = z + \delta \mathbb{E}(f(S(\cdot); \omega, \alpha, \beta), f_\epsilon)$ is value of unemployment with $S(\cdot)$ match surplus function

# DM: From Imperfect Competition to Search

- DM applicability goes beyond our class of imperfectly competitive models

- Consider a standard wage equation of search models inspired by output technology of Bagger & al. (2014)

$$w_{n,t}(d) = \omega \gamma_t(d)^{\alpha} {H_{n,1}}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta) U(H_{n,1})$$

  - ▷ $H_{n,1}$ is HK (known and time-invariant for simplicity)
  - ▷ $\omega$ is workers' bargaining weight (known)
  - ▷ $\gamma_t(d)$ is firm/job productivity (unknown time-varying firm effect)
  - ▷ $\alpha$ and $\beta$ are parameters (firm-invariant for simplicity)
  - ▷ $U(H_{n,1}) = z + \delta \mathbb{E}(f(S(\cdot); \omega, \alpha, \beta), f_{\epsilon})$ is value of unemployment with $S(\cdot)$ match surplus function

# DM: From Imperfect Competition to Search

- DM applicability goes beyond our class of imperfectly competitive models

- Consider a standard wage equation of search models inspired by output technology of Bagger & al. (2014)

$$w_{n,t}(d) = \omega \gamma_t(d)^{\alpha} {H_{n,1}}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta) U(H_{n,1})$$

  - ▷ $H_{n,1}$ is HK (known and time-invariant for simplicity)
  - ▷ $\omega$ is workers' bargaining weight (known)
  - ▷ $\gamma_t(d)$ is firm/job productivity (unknown time-varying firm effect)
  - ▷ $\alpha$ and $\beta$ are parameters (firm-invariant for simplicity)
  - ▷ $U(H_{n,1}) = z + \delta \mathbb{E}(f(S(\cdot); \omega, \alpha, \beta), f_{\epsilon})$ is value of unemployment with $S(\cdot)$ match surplus function

# DM: From Imperfect Competition to Search

- Wage equation

$$w_{n,t}(d) = \omega L_t(d)^{\alpha} H_{n,1}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta)U(H_{n,1})$$

- Intuition: this is a scale/location model, hence can be identified using DM
- Assume w/o loss that $H_{n,1}$ can take value one
- Consider period $t$ and firm $d$; assume $\gamma_{d,t}^{\alpha} = c$ with $c \neq 1$ known
- Then, $\omega \gamma_{d,t}^{\alpha} H_{n,1}^{\beta}$ known at $H_{n,1} = 1$ (DM scale normalisation)
- Assume $(1-\omega)(1-\delta)U(H_{n,1})$ known at $H_{n,1} = 1$ (DM location normalisation)
- Using DM, we identify $\alpha$, $\beta$, $\gamma_{d,t}$, and $U(H_{n,1})$
- Compare largest wage of workers with same $H_{n,1}$ employed by $d, d'$ and identify $\gamma_{d',t}^{\alpha}$, and thus $\gamma_{d',t}$
- Move to period $\tau \neq t$ and assume $\gamma_{d,\tau} = r$ with $r$ known for some $d$
- Compare largest wage of workers w/ same $H_{n,1}$ employed by $d, d'$ and identify $\gamma_{d',\tau}^{\alpha}$, and thus $\gamma_{d',\tau}$ [Back]

# DM: From Imperfect Competition to Search

- Wage equation

$$w_{n,t}(d) = \omega L_t(d)^{\alpha} H_{n,1}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta)U(H_{n,1})$$

- Intuition: this is a scale/location model, hence can be identified using DM
- Assume w/o loss that $H_{n,1}$ can take value one
- Consider period $t$ and firm $d$; assume $\gamma_{d,t}^{\alpha} = c$ with $c \neq 1$ known
- Then, $\omega \gamma_{d,t}^{\alpha} H_{n,1}^{\beta}$ known at $H_{n,1} = 1$ (DM scale normalisation)
- Assume $(1-\omega)(1-\delta)U(H_{n,1})$ known at $H_{n,1} = 1$ (DM location normalisation)
- Using DM, we identify $\alpha$, $\beta$, $\gamma_{d,t}$, and $U(H_{n,1})$
- Compare largest wage of workers with same $H_{n,1}$ employed by $d, d'$ and identify $\gamma_{d',t}^{\alpha}$, and thus $\gamma_{d',t}$
- Move to period $\tau \neq t$ and assume $\gamma_{d,\tau} = r$ with $r$ known for some $d$
- Compare largest wage of workers w/ same $H_{n,1}$ employed by $d, d'$ and identify $\gamma_{d',\tau}^{\alpha}$, and thus $\gamma_{d',\tau}$ [Back]

# DM: From Imperfect Competition to Search

- Wage equation

$$w_{n,t}(d) = \omega L_t(d)^\alpha H_{n,1}^\beta \varepsilon_{n,t}(d) + (1-\omega)(1-\delta)U(H_{n,1})$$

- Intuition: this is a scale/location model, hence can be identified using DM
- Assume w/o loss that $H_{n,1}$ can take value one
- Consider period $t$ and firm $d$; assume $\gamma_{d,t}^\alpha = c$ with $c \neq 1$ known
- Then, $\omega\gamma_{d,t}^\alpha H_{n,1}^\beta$ known at $H_{n,1} = 1$ (DM scale normalisation)
- Assume $(1-\omega)(1-\delta)U(H_{n,1})$ known at $H_{n,1} = 1$ (DM location normalisation)
- Using DM, we identify $\alpha$, $\beta$, $\gamma_{d,t}$, and $U(H_{n,1})$
- Compare largest wage of workers with same $H_{n,1}$ employed by $d, d'$ and identify $\gamma_{d',t}^\alpha$, and thus $\gamma_{d',t}$
- Move to period $\tau \neq t$ and assume $\gamma_{d,\tau} = r$ with $r$ known for some $d$
- Compare largest wage of workers w/ same $H_{n,1}$ employed by $d, d'$ and identify $\gamma_{d',\tau}^\alpha$, and thus $\gamma_{d',\tau}$

[Back]

# DM: From Imperfect Competition to Search

- Wage equation

$$w_{n,t}(d) = \omega L_t(d)^{\alpha} H_{n,1}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta)U(H_{n,1})$$

- Intuition: this is a scale/location model, hence can be identified using DM
- Assume w/o loss that $H_{n,1}$ can take value one
- Consider period $t$ and firm $d$; assume $\gamma_{d,t}^{\alpha} = c$ with $c \neq 1$ known
- Then, $\omega \gamma_{d,t}^{\alpha} H_{n,1}^{\beta}$ known at $H_{n,1} = 1$ (DM scale normalisation)
- Assume $(1-\omega)(1-\delta)U(H_{n,1})$ known at $H_{n,1} = 1$ (DM location normalisation)
- Using DM, we identify $\alpha$, $\beta$, $\gamma_{d,t}$, and $U(H_{n,1})$
- Compare largest wage of workers with same $H_{n,1}$ employed by $d, d'$ and identify $\gamma_{d',t}^{\alpha}$, and thus $\gamma_{d',t}$
- Move to period $\tau \neq t$ and assume $\gamma_{d,\tau} = r$ with $r$ known for some $d$
- Compare largest wage of workers w/ same $H_{n,1}$ employed by $d, d'$ and identify $\gamma_{d',\tau}^{\alpha}$, and thus $\gamma_{d',\tau}$ [Back]

# DM: From Imperfect Competition to Search

- Wage equation

$$w_{n,t}(d) = \omega L_t(d)^{\alpha} H_{n,1}^{\beta} \varepsilon_{n,t}(d) + (1-\omega)(1-\delta)U(H_{n,1})$$

- Intuition: this is a scale/location model, hence can be identified using DM
- Assume w/o loss that $H_{n,1}$ can take value one
- Consider period $t$ and firm $d$; assume $\gamma_{d,t}^{\alpha} = c$ with $c \neq 1$ known
- Then, $\omega \gamma_{d,t}^{\alpha} H_{n,1}^{\beta}$ known at $H_{n,1} = 1$ (DM scale normalisation)
- Assume $(1-\omega)(1-\delta)U(H_{n,1})$ known at $H_{n,1} = 1$ (DM location normalisation)
- Using DM, we identify $\alpha$, $\beta$, $\gamma_{d,t}$, and $U(H_{n,1})$
- Compare largest wage of workers with same $H_{n,1}$ employed by $d, d'$ and identify $\gamma_{d',t}^{\alpha}$, and thus $\gamma_{d',t}$
- Move to period $\tau \neq t$ and assume $\gamma_{d,\tau} = r$ with $r$ known for some $d$
- Compare largest wage of workers w/ same $H_{n,1}$ employed by $d, d'$ and identify $\gamma_{d',\tau}^{\alpha}$, and thus $\gamma_{d',\tau}$ [Back]

# Normalizations

- Location normalization to identify deterministic wage $\varphi$: $\varphi(\cdot)$ known for each firm at one state
- Empirics: we do not need this normalization
  - ▷ $\varphi(\cdot)$ is parameterized with time-invariant parameters
  - ▷ Normalization would correspond to imposing intercept $e + \beta_0(d) = 0$
  - ▷ By exploiting longitudinal dimension, we can identify intercept
  - ▷ Moreover, we can separately identify $e$ and $\beta_0(d)$ by exploiting job mobility as in the AKM
- Location normalization to identify output technology $y(\cdot)$: $y(\cdot)$ known for each firm at one state
- Empirics: do not need this normalization by parameterizing $\varphi(\cdot)$ with time-invariant parameters

[Back]

# Normalizations

- Location normalization to identify deterministic wage $\varphi$: $\varphi(\cdot)$ known for each firm at one state
- Empirics: we do not need this normalization
  - ▷ $\varphi(\cdot)$ is parameterized with time-invariant parameters
  - ▷ Normalization would correspond to imposing intercept $e + \beta_0(d) = 0$
  - ▷ By exploiting longitudinal dimension, we can identify intercept
  - ▷ Moreover, we can separately identify $e$ and $\beta_0(d)$ by exploiting job mobility as in the AKM
- Location normalization to identify output technology $y(\cdot)$: $y(\cdot)$ known for each firm at one state
- Empirics: do not need this normalization by parameterizing $\varphi(\cdot)$ with time-invariant parameters [Back]

# Continuous Efficiency Types $e_n$

- Finite $\mathcal{E}$ with known cardinality ensures identification of wage mixture
- Readily extends to known upper bound $E^*$ on $|\mathcal{E}|$ since Bruni and Koch (1985) accommodate zero weights
- $e_n$ is continuous/multidim.: wage mixture is *continuous* mixture of continuous Gaussian mixtures
  - ▷ Simplify model by removing selection on $\epsilon_{n,t}$: $\epsilon_{n,t}$ conditional on $D_{n,t}$ is distributed as $\epsilon_{n,t}$, e.g. Normal
  - ▷ Wage mixture is continuous mixture of Normals, identified by Bruni and Koch (1985)

# Continuous Efficiency Types $e_n$

- Finite $\mathcal{E}$ with known cardinality ensures identification of wage mixture
- Readily extends to known upper bound $E^*$ on $|\mathcal{E}|$ since Bruni and Koch (1985) accommodate zero weights
- $e_n$ is continuous/multidim.: wage mixture is *continuous* mixture of continuous Gaussian mixtures
  - ▷ Simplify model by removing selection on $\epsilon_{n,t}$: $\epsilon_{n,t}$ conditional on $D_{n,t}$ is distributed as $\epsilon_{n,t}$, e.g. Normal
  - ▷ Wage mixture is continuous mixture of Normals, identified by Bruni and Koch (1985)

# Continuous Signals $a_{n,t}$

- As for $\mathcal{E}$, wage mixture identification can be readily adapted to known upper bound $A^*$ on $|\mathcal{A}|$
- Also covers continuous/multidimensional $a_{n,t}$ with no selection on $\epsilon_{n,t}$
- $|\mathcal{A}| = 2$ simplifies identification of learning process
  - ▷ Signal distribution is *binomial* mixture over $\theta_n$ identified by Blischke (1964; 1978)
- Extends to other cardinalities or continuous/multidim. $a_{n,t}$ if signal mixture over $\theta_n$ remains identifiable
  - ▷ E.g. if $a_{n,t}$ is continuous/multivariate Gaussian mixture conditional on $\theta_n$
  - ▷ Signal is finite mixture of continuous/multiv. Gaussian mixtures, identified by Bruni and Koch (1985)

# Continuous Signals $a_{n,t}$

- As for $\mathcal{E}$, wage mixture identification can be readily adapted to known upper bound $A^*$ on $|\mathcal{A}|$
- Also covers continuous/multidimensional $a_{n,t}$ with no selection on $\epsilon_{n,t}$
- $|\mathcal{A}| = 2$ simplifies identification of learning process
  - ▷ Signal distribution is *binomial* mixture over $\theta_n$ identified by Blischke (1964; 1978)
- Extends to other cardinalities or continuous/multidim. $a_{n,t}$ if signal mixture over $\theta_n$ remains identifiable
  - ▷ E.g. if $a_{n,t}$ is continuous/multivariate Gaussian mixture conditional on $\theta_n$
  - ▷ Signal is finite mixture of continuous/multiv. Gaussian mixtures, identified by Bruni and Koch (1985)

# Continuous Signals $a_{n,t}$

- As for $\mathcal{E}$, wage mixture identification can be readily adapted to known upper bound $A^*$ on $|\mathcal{A}|$
- Also covers continuous/multidimensional $a_{n,t}$ with no selection on $\epsilon_{n,t}$
- $|\mathcal{A}| = 2$ simplifies identification of learning process
  - ▷ Signal distribution is *binomial* mixture over $\theta_n$ identified by Blischke (1964; 1978)
- Extends to other cardinalities or continuous/multidim. $a_{n,t}$ if signal mixture over $\theta_n$ remains identifiable
  - ▷ E.g. if $a_{n,t}$ is continuous/multivariate Gaussian mixture conditional on $\theta_n$
  - ▷ Signal is finite mixture of continuous/multiv. Gaussian mixtures, identified by Bruni and Koch (1985)

# Continuous Ability $\theta_n$

- $|\Theta| = 2$ simplifies identification of learning process
  - ▷ Signal distribution is binomial mixture over $\theta_n$ with *two* components
  - ▷ Identified by Blischke (1964; 1978) if workers stay in a job for at least three periods
- Extends to other finite cardinalities provided that we increase the required length of job retention
- Extends to continuous/multidimensional $\theta_n$ if signal mixture over $\theta_n$ remains identifiable
  - ▷ E.g. if $a_{n,t}$ is continuous/multivariate Normal conditional on $\theta_n$
  - ▷ Signal is continuous mixture of multivariate Normals, identified by Bruni and Koch (1985) [Back]

# Continuous Ability $\theta_n$

- $|\Theta| = 2$ simplifies identification of learning process
  - ▷ Signal distribution is binomial mixture over $\theta_n$ with *two* components
  - ▷ Identified by Blischke (1964; 1978) if workers stay in a job for at least three periods
- Extends to other finite cardinalities provided that we increase the required length of job retention
- Extends to continuous/multidimensional $\theta_n$ if signal mixture over $\theta_n$ remains identifiable
  - ▷ E.g. if $a_{n,t}$ is continuous/multivariate Normal conditional on $\theta_n$
  - ▷ Signal is continuous mixture of multivariate Normals, identified by Bruni and Koch (1985) [Back]

# Continuous Ability $\theta_n$

- $|\Theta| = 2$ simplifies identification of learning process
  - ▷ Signal distribution is binomial mixture over $\theta_n$ with *two* components
  - ▷ Identified by Blischke (1964; 1978) if workers stay in a job for at least three periods
- Extends to other finite cardinalities provided that we increase the required length of job retention
- Extends to continuous/multidimensional $\theta_n$ if signal mixture over $\theta_n$ remains identifiable
  - ▷ E.g. if $a_{n,t}$ is continuous/multivariate Normal conditional on $\theta_n$
  - ▷ Signal is continuous mixture of multivariate Normals, identified by Bruni and Koch (1985) [Back]

# Which Tail of the Wage Distribution?

- Wage of job $d$ is 2nd-best firm output $y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)$ plus comp differential $\Psi(d, d', s_{n,t}(e))$
- Thus, by letting the wage of firm $d$ go to $+\infty$, while the state variables $s_{n,t}(e)$ remain fixed
- We effectively push productivity shock of 2nd-best firm $\epsilon_{n,t}(d', e)$ to $+\infty$
- We potentially displace firm $d$ from being the top choice for worker $n$
- We make the probability of choosing firm $d$ going to zero
- We solve the issue and preserve firm ranking by sending wages to $-\infty$ (left tail of wage distribution) [Back]

# Which Tail of the Wage Distribution?

- Wage of job $d$ is 2nd-best firm output $y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)$ plus comp differential $\Psi(d, d', s_{n,t}(e))$
- Thus, by letting the wage of firm $d$ go to $+\infty$, while the state variables $s_{n,t}(e)$ remain fixed
- We effectively push productivity shock of 2nd-best firm $\epsilon_{n,t}(d', e)$ to $+\infty$
- We potentially displace firm $d$ from being the top choice for worker $n$
- We make the probability of choosing firm $d$ going to zero
- We solve the issue and preserve firm ranking by sending wages to $-\infty$ (left tail of wage distribution) [Back]

# Which Tail of the Wage Distribution?

- Wage of job $d$ is 2nd-best firm output $y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)$ plus comp differential $\Psi(d, d', s_{n,t}(e))$
- Thus, by letting the wage of firm $d$ go to $+\infty$, while the state variables $s_{n,t}(e)$ remain fixed
- We effectively push productivity shock of 2nd-best firm $\epsilon_{n,t}(d', e)$ to $+\infty$
- We potentially displace firm $d$ from being the top choice for worker $n$
- We make the probability of choosing firm $d$ going to zero
- We solve the issue and preserve firm ranking by sending wages to $-\infty$ (left tail of wage distribution) [Back]

# Identification of Information Technology

- Concatenate wage mixture weights across periods to identify prior and conditional signal distribution

$$p_1(H_{n,1}, e_n) \quad \text{and} \quad \Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$$

- How? In four steps
  - ▷ Identify $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ from wage mixture weights at $t$ and $t+1$ w/o further assumptions
  - ▷ Represent $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ as a mixture over $\theta_n$
  - ▷ Show identification of signal mixture under various sets of conditions
    - ○ E.g. $\mathcal{A} := \{\underline{a}, \bar{a}\}$ and $\Theta := \{\underline{\theta}, \bar{\theta}\}$: binomial mixture with 2 components (Blischke, 1964; 1978)
    - ○ Can generalise to continuous mixtures, e.g. using Bruni and Koch (1985)
  - ▷ Identify $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$ from signal mixture weights and components
- By combining $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$, we compute $P_{n,t}$ and its distribution [Back]

# Identification of Information Technology

- Concatenate wage mixture weights across periods to identify prior and conditional signal distribution

$$p_1(H_{n,1}, e_n) \quad \text{and} \quad \Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$$

- How? In four steps
  - ▷ Identify $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ from wage mixture weights at $t$ and $t+1$ w/o further assumptions
  - ▷ Represent $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ as a mixture over $\theta_n$
  - ▷ Show identification of signal mixture under various sets of conditions
    - ○ E.g. $\mathcal{A} := \{\underline{a}, \bar{a}\}$ and $\Theta := \{\underline{\theta}, \bar{\theta}\}$: binomial mixture with 2 components (Blischke, 1964; 1978)
    - ○ Can generalise to continuous mixtures, e.g. using Bruni and Koch (1985)
  - ▷ Identify $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$ from signal mixture weights and components
- By combining $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$, we compute $P_{n,t}$ and its distribution [Back]

# Identification of Information Technology

- Concatenate wage mixture weights across periods to identify prior and conditional signal distribution

$$p_1(H_{n,1}, e_n) \quad \text{and} \quad \Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$$

- How? In four steps
  - ▷ Identify $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ from wage mixture weights at $t$ and $t+1$ w/o further assumptions
  - ▷ Represent $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ as a mixture over $\theta_n$
  - ▷ Show identification of signal mixture under various sets of conditions
    - ○ E.g. $\mathcal{A} := \{\underline{a}, \bar{a}\}$ and $\Theta := \{\underline{\theta}, \bar{\theta}\}$: binomial mixture with 2 components (Blischke, 1964; 1978)
    - ○ Can generalise to continuous mixtures, e.g. using Bruni and Koch (1985)
  - ▷ Identify $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$ from signal mixture weights and components
- By combining $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$, we compute $P_{n,t}$ and its distribution [Back]

# Identification of Information Technology

- Concatenate wage mixture weights across periods to identify prior and conditional signal distribution

$$p_1(H_{n,1}, e_n) \quad \text{and} \quad \Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$$

- How? In four steps
  - ▷ Identify $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ from wage mixture weights at $t$ and $t+1$ w/o further assumptions
  - ▷ Represent $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ as a mixture over $\theta_n$
  - ▷ Show identification of signal mixture under various sets of conditions
    - ○ E.g. $\mathcal{A} := \{\underline{a}, \bar{a}\}$ and $\Theta := \{\underline{\theta}, \bar{\theta}\}$: binomial mixture with 2 components (Blischke, 1964; 1978)
    - ○ Can generalise to continuous mixtures, e.g. using Bruni and Koch (1985)
  - ▷ Identify $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$ from signal mixture weights and components
- By combining $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$, we compute $P_{n,t}$ and its distribution [Back]

# Identification of Information Technology

- Concatenate wage mixture weights across periods to identify prior and conditional signal distribution

$$p_1(H_{n,1}, e_n) \quad \text{and} \quad \Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$$

- How? In four steps
  - ▷ Identify $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ from wage mixture weights at $t$ and $t+1$ w/o further assumptions
  - ▷ Represent $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ as a mixture over $\theta_n$
  - ▷ Show identification of signal mixture under various sets of conditions
    - ○ E.g. $\mathcal{A} := \{\underline{a}, \bar{a}\}$ and $\Theta := \{\underline{\theta}, \bar{\theta}\}$: binomial mixture with 2 components (Blischke, 1964; 1978)
    - ○ Can generalise to continuous mixtures, e.g. using Bruni and Koch (1985)
  - ▷ Identify $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$ from signal mixture weights and components
- By combining $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$, we compute $P_{n,t}$ and its distribution [Back]

# Identification of Information Technology

- Concatenate wage mixture weights across periods to identify prior and conditional signal distribution

$$p_1(H_{n,1}, e_n) \quad \text{and} \quad \Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$$

- How? In four steps
  - ▷ Identify $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ from wage mixture weights at $t$ and $t+1$ w/o further assumptions
  - ▷ Represent $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ as a mixture over $\theta_n$
  - ▷ Show identification of signal mixture under various sets of conditions
    - ○ E.g. $\mathcal{A} := \{\underline{a}, \bar{a}\}$ and $\Theta := \{\underline{\theta}, \bar{\theta}\}$: binomial mixture with 2 components (Blischke, 1964; 1978)
    - ○ Can generalise to continuous mixtures, e.g. using Bruni and Koch (1985)
  - ▷ Identify $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$ from signal mixture weights and components
- By combining $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$, we compute $P_{n,t}$ and its distribution [Back]

# Identification of Information Technology

- Concatenate wage mixture weights across periods to identify prior and conditional signal distribution

$$p_1(H_{n,1}, e_n) \quad \text{and} \quad \Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$$

- How? In four steps
  - ▷ Identify $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ from wage mixture weights at $t$ and $t+1$ w/o further assumptions
  - ▷ Represent $\Pr(a_n^t \mid H_{n,1}, D_n^t, e_n)$ as a mixture over $\theta_n$
  - ▷ Show identification of signal mixture under various sets of conditions
    - ○ E.g. $\mathcal{A} := \{\underline{a}, \bar{a}\}$ and $\Theta := \{\underline{\theta}, \bar{\theta}\}$: binomial mixture with 2 components (Blischke, 1964; 1978)
    - ○ Can generalise to continuous mixtures, e.g. using Bruni and Koch (1985)
  - ▷ Identify $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$ from signal mixture weights and components
- By combining $p_1(H_{n,1}, e_n)$ and $\Pr(a_{n,t-1} \mid H_{n,1}, D_{n,t-1}, e_n, \theta_n)$, we compute $P_{n,t}$ and its distribution [Back]

# Identification of Deterministic Wage

- Having identified distribution of $(D_{n,t}, w_{n,t})$ given $s_{n,t}$, we use DM to identify $\varphi(\cdot) := y(\cdot) + \Psi(\cdot)$
- However, due to second-price auction-like mechanism, we must send wages to $-\infty$ rather than $+\infty$
- Wage of job $d$ is 2nd-best firm output $y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)$ plus comp differential $\Psi(d, d', s_{n,t}(e))$
- Thus, by letting the wage of job $d$ go to $+\infty$, while the state variables $s_{n,t}(e)$ remain fixed
- We effectively push productivity shock of 2nd-best firm $\epsilon_{n,t}(d', e)$ to $+\infty$
- We potentially displace job $d$ from being the top choice for worker $n$
- We make the probability of choosing job $d$ going to zero
- Solve this by sending wages to $-\infty$ (left tail of the wage distribution) [Back]

# Identification of Deterministic Wage

- Having identified distribution of $(D_{n,t}, w_{n,t})$ given $s_{n,t}$, we use DM to identify $\varphi(\cdot) := y(\cdot) + \Psi(\cdot)$
- However, due to second-price auction-like mechanism, we must send wages to $-\infty$ rather than $+\infty$
- Wage of job $d$ is 2nd-best firm output $y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)$ plus comp differential $\Psi(d, d', s_{n,t}(e))$
- Thus, by letting the wage of job $d$ go to $+\infty$, while the state variables $s_{n,t}(e)$ remain fixed
- We effectively push productivity shock of 2nd-best firm $\epsilon_{n,t}(d', e)$ to $+\infty$
- We potentially displace job $d$ from being the top choice for worker $n$
- We make the probability of choosing job $d$ going to zero
- Solve this by sending wages to $-\infty$ (left tail of the wage distribution) [Back]

# Identification of Deterministic Wage

- Having identified distribution of $(D_{n,t}, w_{n,t})$ given $s_{n,t}$, we use DM to identify $\varphi(\cdot) := y(\cdot) + \Psi(\cdot)$
- However, due to second-price auction-like mechanism, we must send wages to $-\infty$ rather than $+\infty$
- Wage of job $d$ is 2nd-best firm output $y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)$ plus comp differential $\Psi(d, d', s_{n,t}(e))$
- Thus, by letting the wage of job $d$ go to $+\infty$, while the state variables $s_{n,t}(e)$ remain fixed
- We effectively push productivity shock of 2nd-best firm $\epsilon_{n,t}(d', e)$ to $+\infty$
- We potentially displace job $d$ from being the top choice for worker $n$
- We make the probability of choosing job $d$ going to zero
- Solve this by sending wages to $-\infty$ (left tail of the wage distribution) [Back]

# Identification of Deterministic Wage

- Having identified distribution of $(D_{n,t}, w_{n,t})$ given $s_{n,t}$, we use DM to identify $\varphi(\cdot) := y(\cdot) + \Psi(\cdot)$
- However, due to second-price auction-like mechanism, we must send wages to $-\infty$ rather than $+\infty$
- Wage of job $d$ is 2nd-best firm output $y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)$ plus comp differential $\Psi(d, d', s_{n,t}(e))$
- Thus, by letting the wage of job $d$ go to $+\infty$, while the state variables $s_{n,t}(e)$ remain fixed
- We effectively push productivity shock of 2nd-best firm $\epsilon_{n,t}(d', e)$ to $+\infty$
- We potentially displace job $d$ from being the top choice for worker $n$
- We make the probability of choosing job $d$ going to zero
- Solve this by sending wages to $-\infty$ (left tail of the wage distribution) [Back]

# Identification of Deterministic Wage

- Having identified distribution of $(D_{n,t}, w_{n,t})$ given $s_{n,t}$, we use DM to identify $\varphi(\cdot) := y(\cdot) + \Psi(\cdot)$
- However, due to second-price auction-like mechanism, we must send wages to $-\infty$ rather than $+\infty$
- Wage of job $d$ is 2nd-best firm output $y(d', s_{n,t}(e)) + \epsilon_{n,t}(d', e)$ plus comp differential $\Psi(d, d', s_{n,t}(e))$
- Thus, by letting the wage of job $d$ go to $+\infty$, while the state variables $s_{n,t}(e)$ remain fixed
- We effectively push productivity shock of 2nd-best firm $\epsilon_{n,t}(d', e)$ to $+\infty$
- We potentially displace job $d$ from being the top choice for worker $n$
- We make the probability of choosing job $d$ going to zero
- Solve this by sending wages to $-\infty$ (left tail of the wage distribution) [Back]

# Identification of Output and HK Technology

- Equilibrium is efficient
- Market-wide equilibrium allocation problem reduces to single-agent (planner) dynamic decision problem
- We have identified CCPs and distribution of productivity shocks
- Therefore, $y(\cdot)$ can be identified following Magnac & Thesmar (2002)
- Under usual normalisations: $y(\cdot)$ is known at one firm for each state
- Can be extended to multi-job case where equilibrium is typically inefficient
  - ▷ Market-wide equilibrium allocation solves pseudo-planning
  - ▷ Namely, problem of maximising match surplus for each firm [Back]

# Identification of Output and HK Technology

- Equilibrium is efficient
- Market-wide equilibrium allocation problem reduces to single-agent (planner) dynamic decision problem
- We have identified CCPs and distribution of productivity shocks
- Therefore, $y(\cdot)$ can be identified following Magnac & Thesmar (2002)
- Under usual normalisations: $y(\cdot)$ is known at one firm for each state
- Can be extended to multi-job case where equilibrium is typically inefficient
  - Market-wide equilibrium allocation solves pseudo-planning
  - Namely, problem of maximising match surplus for each firm [Back]

# Identification of Output and HK Technology

- Equilibrium is efficient
- Market-wide equilibrium allocation problem reduces to single-agent (planner) dynamic decision problem
- We have identified CCPs and distribution of productivity shocks
- Therefore, $y(\cdot)$ can be identified following Magnac & Thesmar (2002)
- Under usual normalisations: $y(\cdot)$ is known at one firm for each state
- Can be extended to multi-job case where equilibrium is typically inefficient
  - ▷ Market-wide equilibrium allocation solves pseudo-planning
  - ▷ Namely, problem of maximising match surplus for each firm [Back]

# Identification of CCPs

- Recall that wage mixture weights essentially determine distribution of $(H_{n,1}, D_n^{t-1}, e_n, a_n^{t-1})$ in each period
- This, in turn, governs distribution of state variables $s_{n,t}$ and subsequently occupation choices
- Therefore, by appropriately combining weights across periods, we recover CCPs
- Unlike common approaches, here CCPs are identified from the continuous part (wage distribution) [Back]

# Identification of CCPs

- Recall that wage mixture weights essentially determine distribution of $(H_{n,1}, D_n^{t-1}, e_n, a_n^{t-1})$ in each period
- This, in turn, governs distribution of state variables $s_{n,t}$ and subsequently occupation choices
- Therefore, by appropriately combining weights across periods, we recover CCPs
- Unlike common approaches, here CCPs are identified from the continuous part (wage distribution) [Back]

# Identification of CCPs

- Recall that wage mixture weights essentially determine distribution of $(H_{n,1}, D_n^{t-1}, e_n, a_n^{t-1})$ in each period
- This, in turn, governs distribution of state variables $s_{n,t}$ and subsequently occupation choices
- Therefore, by appropriately combining weights across periods, we recover CCPs
- Unlike common approaches, here CCPs are identified from the continuous part (wage distribution) [Back]

# Equilibrium: Robust MPE

- Refinement condition: non-employing firms indifferent between not employing and employing the worker
- Without refinement condition: multiplicity of qualitatively *similar* MPE (*same* on-path outcomes)
  - ▷ *Same* ranking of firms in terms of the EPDV of offered wages
  - ▷ Differ in terms of the wages offered by the non-employing firms
  - ▷ Non-employing firms can offer any wage up to indifference
- Refinement solves this *trivial* multiplicity of equilibria and selects one equilibrium
- In a way that is standard in the literature on trembling-hand perfect equilibrium (Selten, 1975) [Back]

# Equilibrium: Robust MPE

- Refinement condition: non-employing firms indifferent between not employing and employing the worker
- Without refinement condition: multiplicity of qualitatively *similar* MPE (*same* on-path outcomes)
  - ▷ *Same* ranking of firms in terms of the EPDV of offered wages
  - ▷ Differ in terms of the wages offered by the non-employing firms
  - ▷ Non-employing firms can offer any wage up to indifference
- Refinement solves this *trivial* multiplicity of equilibria and selects one equilibrium
- In a way that is standard in the literature on trembling-hand perfect equilibrium (Selten, 1975) [Back]

# Equilibrium: Robust MPE

- Refinement condition: non-employing firms indifferent between not employing and employing the worker
- Without refinement condition: multiplicity of qualitatively *similar* MPE (*same* on-path outcomes)
  - ▷ *Same* ranking of firms in terms of the EPDV of offered wages
  - ▷ Differ in terms of the wages offered by the non-employing firms
  - ▷ Non-employing firms can offer any wage up to indifference
- Refinement solves this *trivial* multiplicity of equilibria and selects one equilibrium
- In a way that is standard in the literature on trembling-hand perfect equilibrium (Selten, 1975) [Back]

# Equilibrium: Robust MPE

- Refinement condition: non-employing firms indifferent between not employing and employing the worker
- Without refinement condition: multiplicity of qualitatively *similar* MPE (*same* on-path outcomes)
  - ▷ *Same* ranking of firms in terms of the EPDV of offered wages
  - ▷ Differ in terms of the wages offered by the non-employing firms
  - ▷ Non-employing firms can offer any wage up to indifference
- Refinement solves this *trivial* multiplicity of equilibria and selects one equilibrium
- In a way that is standard in the literature on trembling-hand perfect equilibrium (Selten, 1975) [Back]

# Equilibrium: Worker Bellman Equation

$$\tilde{W}(s_{n,t}(e), \epsilon_{n,t}(e), \{w_{d,n,t}(e)\}_{d\in\mathcal{D}}) = \max_{\{l_{d,n,t}(e)\}_{d\in\mathcal{D}}} \sum_{d\in\mathcal{D}} l_{d,n,t}(e) \times \Big[ w_{d,n,t}(e)$$
$$+ \delta[1-\eta(\kappa_{n,t}, d)] \int_{\epsilon_{n,t+1}(e)} \mathbb{E}\Big( \tilde{W}(s_{n,t+1}(e), \epsilon_{n,t+1}(e), \{w_{d,n,t+1}(e)\}_{d\in\mathcal{D}}) \mid s_{n,t}(e), d\Big) dG_e \Big]$$

# Equilibrium: Firm Bellman Equation

$$\Pi_d(s_{n,t}(e), \epsilon_{n,t}(e)) = \max_{w_{d,n,t}(e)} \Big( l_{d,n,t}(e) \times \Big[ y(d, s_{n,t}(e)) + \epsilon_{n,t}(d, e) - w_{d,n,t}(e)$$

$$+ \delta[1 - \eta(\kappa_{n,t}, d)] \int_{\epsilon_{n,t+1}(e)} \mathbb{E}\Big( \Pi_d(s_{n,t+1}(e), \epsilon_{n,t+1}(e)) \mid s_{n,t}(e), d \Big) dG_e \Big]$$

$$+ \sum_{d' \in \mathcal{D} \setminus \{d\}} l_{d',n,t}(e) \times \Big\{ \delta[1 - \eta(\kappa_{n,t}, d')] \int_{\epsilon_{n,t+1}(e)} \mathbb{E}\Big( \Pi_d(s_{n,t+1}(e), \epsilon_{n,t+1}(e)) \mid s_{n,t}(e), d' \Big) dG_e \Big\} \Big)$$

# Equilibrium: Refinement

- To ensure uniqueness, suppose firm $d'$ employs worker $n$ at state $(s_{n,t}(e), \epsilon_{n,t}(e))$
- Then the offer by each non-employing firm $d \neq d'$ must make $d$ indifferent between
  - ▷ Not employing $n$ (LHS)
  - ▷ Employing $n$ (RHS)

$$\delta[1 - \eta(\kappa_{n,t}, d')] \int_{\epsilon_{n,t+1}(e)} \mathbb{E}\Pi_d(\cdot | s_{n,t}(e), d') dG_e$$
$$= \max_{w_{d,n,t}(e)} \left\{ y(d, s_{n,t}(e)) + \epsilon_{n,t}(d,e) - w_{d,n,t}(e) + \delta[1 - \eta(\kappa_{n,t}, d)] \int_{\epsilon_{n,t+1}(e)} \mathbb{E}\Pi_d(\cdot | s_{n,t}(e), d) dG_e \right\}$$

[Back]

# Mixture: Technical Condition for Identification

- The wage mixture $f$ satisfies the clusterability condition if

$$\inf_{i \neq j} \rho(\gamma_i, \gamma_j) > (4 + \xi_\Lambda)\eta(\Lambda) \quad \text{for some } \xi_\Lambda > 0$$

- $\gamma_i, \gamma_j$: mixture components
- $\rho$: metric, e.g. Hellinger and total variation
- $\Gamma$: mixing measure of $f$, i.e., function mapping mixture components to weights
- $\eta(\Gamma)$: measure of asymptotic diameter of approximating mixture measures
- As $L \to \infty$, $\gamma_k$ must be separated by gap proportional to diameter of approximating mixture measures [Back]

# Wage Mixture Identification: Gaussian Case

- Continuous r.v. $W$ with PDF $f_W(\cdot)$; $D$ is compact subset of $\mathbb{R} \times \mathbb{R}^+$; $g(\cdot; \mu, \sigma^2)$ is Normal pdf
- Claim: if $D$ large enough, there exists probability measure $\pi(\cdot)$ on $D$ such that

$$f_W(w) \approx \int_D g(w; \mu, \sigma^2) d\pi(\mu, \sigma^2) \quad \text{for each } w \in \mathbb{R}$$

Proof.

Let $\mathbb{P}_W(\cdot)$ be probability measure associated with $f_W(\cdot)$. Any $f_W(\cdot)$ can be approximated by convolution of $f_W(\cdot)$ with centered Normal PDF $g(\cdot; 0, s^2)$ for small $s^2$

$$f_W(w) \approx \int_{\mathbb{R}} g(w - \mu; 0, s^2) d\mathbb{P}_W(\mu) = \int_{\mathbb{R}} g(w; \mu, s^2) d\mathbb{P}_W(\mu)$$

Let $D$ be large enough to contain $\mathcal{A}_\tau \times (0, \eta)$, where $\eta \in (0, \infty)$, $\tau$ is small strictly positive number, and $\mathbb{P}_W(\mathcal{A}_\tau) > 1 - \tau$. Then, for small $s^2 \in (0, \eta)$

$$f_W(w) \approx \int_{\mathbb{R}} g(w; \mu, s^2) d\mathbb{P}_W(\mu) \approx \int_{\mathcal{A}_\tau} g(w; \mu, s^2) d\mathbb{P}_W(\mu) = \int_D g(w; \mu, \sigma^2) d\mathbb{P}_W(\mu) \times \mathbb{1}\{\mu \in \mathcal{A}_\tau, \sigma^2 = s^2\}$$

By setting $\pi(\mu, \sigma^2) \equiv \mathbb{P}_W(\mu) \times \mathbb{1}\{\mu \in \mathcal{A}_\tau, \sigma^2 = s^2\}$, we obtain result $\square$

[Back]

# Sketch of Proof of DM

- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then

$$\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \quad \forall x$$

- It implies $\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \,|\, X_n = x)$
- Using survival function $\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \ell_1 S(w - y(1,x))$
- Moreover, if $y(1,\bar{x}) = 0$, $\Pr(D_n = 1, w_n(1) - y(1,x) \geq w \mid X_n = \bar{x}) \sim \ell_1 S(w - y(1,x))$
- Therefore, $y(1,x)$ is identified if

$$\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \Pr(D_n = 1, w_n(1) + u \geq w \,|\, X_n = x) \Rightarrow u = -y(1,x)$$

- Through simple manipulations, the LHS implies $S(w) \sim S(w + u + y(1,x))$
- If wage tails are not excessively thick, $S(w) \sim S(w + u + y(1,x))$ is possible only if $u + y(1,x) = 0$

  ▷ $\mathbb{E}(\exp(\beta\epsilon_n(1))) < +\infty$ for some $\beta > 0$: tails heavier than normal, e.g. Laplace and logistic

[Back] [Lemma]

# Sketch of Proof of DM

- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then

$$\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \quad \forall x$$

- It implies $\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \,|\, X_n = x)$
- Using survival function $\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \ell_1 S(w - y(1,x))$
- Moreover, if $y(1,\bar{x}) = 0$, $\Pr(D_n = 1, w_n(1) - y(1,x) \geq w \mid X_n = \bar{x}) \sim \ell_1 S(w - y(1,x))$
- Therefore, $y(1,x)$ is identified if

$$\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \Pr(D_n = 1, w_n(1) + u \geq w \,|\, X_n = x) \Rightarrow u = -y(1,x)$$

- Through simple manipulations, the LHS implies $S(w) \sim S(w + u + y(1,x))$
- If wage tails are not excessively thick, $S(w) \sim S(w + u + y(1,x))$ is possible only if $u + y(1,x) = 0$
  - ▷ $\mathbb{E}(\exp(\beta\epsilon_n(1))) < +\infty$ for some $\beta > 0$: tails heavier than normal, e.g. Laplace and logistic

[Back] [Lemma]

# Sketch of Proof of DM

- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then

$$\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \quad \forall x$$

- It implies $\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x)$
- Using survival function $\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 S(w - y(1,x))$
- Moreover, if $y(1,\bar{x}) = 0$, $\Pr(D_n = 1, w_n(1) - y(1,x) \geq w \mid X_n = \bar{x}) \sim \ell_1 S(w - y(1,x))$
- Therefore, $y(1,x)$ is identified if

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \Pr(D_n = 1, w_n(1) + u \geq w \mid X_n = x) \Rightarrow u = -y(1,x)$$

- Through simple manipulations, the LHS implies $S(w) \sim S(w + u + y(1,x))$
- If wage tails are not excessively thick, $S(w) \sim S(w + u + y(1,x))$ is possible only if $u + y(1,x) = 0$
    - $\triangleright$ $\mathbb{E}(\exp(\beta\epsilon_n(1))) < +\infty$ for some $\beta > 0$: tails heavier than normal, e.g. Laplace and logistic

[Back] [Lemma]

# Sketch of Proof of DM

- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then

$$\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \quad \forall x$$

- It implies $\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \,|\, X_n = x)$
- Using survival function $\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \ell_1 S(w - y(1,x))$
- Moreover, if $y(1,\bar{x}) = 0$, $\Pr(D_n = 1, w_n(1) - y(1,x) \geq w \mid X_n = \bar{x}) \sim \ell_1 S(w - y(1,x))$
- Therefore, $y(1,x)$ is identified if

$$\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \Pr(D_n = 1, w_n(1) + u \geq w \,|\, X_n = x) \Rightarrow u = -y(1,x)$$

- Through simple manipulations, the LHS implies $S(w) \sim S(w + u + y(1,x))$
- If wage tails are not excessively thick, $S(w) \sim S(w + u + y(1,x))$ is possible only if $u + y(1,x) = 0$
  - ▷ $\mathbb{E}(\exp(\beta\epsilon_n(1))) < +\infty$ for some $\beta > 0$: tails heavier than normal, e.g. Laplace and logistic

[Back] [Lemma]

# Sketch of Proof of DM

- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then

$$\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \quad \forall x$$

- It implies $\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \,|\, X_n = x)$
- Using survival function $\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \ell_1 S(w - y(1, x))$
- Moreover, if $y(1, \bar{x}) = 0$, $\Pr(D_n = 1, w_n(1) - y(1, x) \geq w \mid X_n = \bar{x}) \sim \ell_1 S(w - y(1, x))$
- Therefore, $y(1, x)$ is identified if

$$\Pr(D_n = 1, w_n(1) \geq w \,|\, X_n = x) \sim \Pr(D_n = 1, w_n(1) + u \geq w \,|\, X_n = x) \Rightarrow u = -y(1, x)$$

- Through simple manipulations, the LHS implies $S(w) \sim S(w + u + y(1, x))$
- If wage tails are not excessively thick, $S(w) \sim S(w + u + y(1, x))$ is possible only if $u + y(1, x) = 0$
  - ▷ $\mathbb{E}(\exp(\beta\epsilon_n(1))) < +\infty$ for some $\beta > 0$: tails heavier than normal, e.g. Laplace and logistic

[Back] [Lemma]

# Sketch of Proof of DM

- If $\epsilon_n(1)$ and $\epsilon_n(0)$ "moderately" dependent, then

$$\lim_{w \to +\infty} \Pr(D_n = 1 \mid X_n = x, w_n(1) = w) = \ell_1 > 0 \quad \forall x$$

- It implies $\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 \Pr(w_n(1) \geq w \mid X_n = x)$
- Using survival function $\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \ell_1 S(w - y(1, x))$
- Moreover, if $y(1, \bar{x}) = 0$, $\Pr(D_n = 1, w_n(1) - y(1, x) \geq w \mid X_n = \bar{x}) \sim \ell_1 S(w - y(1, x))$
- Therefore, $y(1, x)$ is identified if

$$\Pr(D_n = 1, w_n(1) \geq w \mid X_n = x) \sim \Pr(D_n = 1, w_n(1) + u \geq w \mid X_n = x) \Rightarrow u = -y(1, x)$$

- Through simple manipulations, the LHS implies $S(w) \sim S(w + u + y(1, x))$
- If wage tails are not excessively thick, $S(w) \sim S(w + u + y(1, x))$ is possible only if $u + y(1, x) = 0$
  - ▷ $\mathbb{E}(\exp(\beta \epsilon_n(1))) < +\infty$ for some $\beta > 0$: tails heavier than normal, e.g. Laplace and logistic

[Back] [Lemma]

# Sketch of Proof of DM

- Let $S$ be survival function of r.v. $Y$ with full support
- Assume $\mathbb{E}\left(\max\{0, Y\}^p\right) < +\infty$ for some $p > 0$
- Let $h$ be a function such that $h(y) \sim y$ and assume $S(y) \sim S\big(\kappa\, h(y)\big)$ for some $\kappa > 0$
- Claim: $\kappa = 1$

Proof.

- By contradiction: suppose $\kappa \neq 1$
- If $\kappa \neq 1$, then $S(y)$ cannot vanish exactly like $y^{-p}$
  - By contradiction: suppose $S(y) \sim y^{-p}$
  - Then, since $h(y) \sim y$, we also have $S\big(\kappa\, h(y)\big) \sim (\kappa\, h(y))^{-p} = \kappa^{-p} y^{-p}$
  - But, by assumption, $S(y) \sim S\big(\kappa\, h(y)\big)$
  - Hence, $y^{-p} \sim S(y) \sim S\big(\kappa\, h(y)\big) \sim \kappa^{-p} y^{-p}$
  - Only way to avoid contradiction here is if $\kappa^{-p} = 1$, meaning $\kappa = 1$
- If $S(y)$ decays either more slowly or more quickly than $y^{-p}$, then $\mathbb{E}\left(\max\{0, Y\}^p\right)$ must be infinite
  - If slower, the integral $\int_0^\infty p\, y^{p-1} S(y)\, dy$ diverges, which implies $\mathbb{E}\left(\max\{0, Y\}^p\right) = +\infty$
  - If faster, the integral $\int_0^\infty p\, y^{p-1} S(y)\, dy$ diverges as well because $S(y) \sim S(\kappa\, y)$ by assumption
- Therefore, $\kappa = 1$

$\square$

[Back]