

A Quantitative Theory of Information, Worker Flows, and Wage Dispersion

Amanda M Michaud*

May 8, 2017

Abstract

Employer learning provides a link between wage and employment dynamics. Workers who are selectively terminated when their low-productivity is revealed subsequently earn lower wages. If learning is asymmetric across employers, randomly separated high-productivity workers are treated similarly when hired from unemployment, but recover as their next employer learns their type. I provide empirical evidence supporting this link, then study whether employer learning is an empirically important factor in wage and employment dynamics. In a calibrated structural model, learning accounts for 78% of wage losses after unemployment, 24% of life-cycle wage growth, and 13% of cross-sectional dispersion observed in data.

JEL Codes: E24, J63, J31

*Indiana University, mail: Wylie Hall 105, 100 S. Woodlawn, Bloomington, IN 47405-7104; e-mail: ammichau@indiana.edu ; phone: (812) 856-1238. v1: March 2010. I thank Bob Becker, Alessandra Fogli, Limor Golan, Jonathan Heathcoate, Henry Hyatt, Rasmus Lentz, Hani Mansour, José Víctor Ríos-Rull, Chris Phelan, Joshua Pinkston, Sam Schulhofer-Wohl, Kjetil Storesletten, and David Wiczer, as well as anonymous referees and Richard Rogerson, for helpful comments and guidance. I acknowledge the Bilinski Foundation for financial support.

What determines the growth of wages over a worker’s career? A body of literature has employed structural models to evaluate whether popular microeconomic theories are each likely to contribute to the empirical wage dynamics in an important way. The most basic theories posit that wages grow as workers become more productive.¹ Such productivity growth may be a consequence of broad human capital growth either by specific investment or as a side effect of experience.² Other theories associate wage growth with employer changes. While this source of wage growth could too reflect higher productivity through better matches, the mere presence of competition precipitating job changes can also generate wage changes in absence of true productivity change.³ In this paper, I study whether an alternative theory, that of asymmetric employer learning, is an empirically important factor in understanding wage and employment dynamics.

The motivation for considering the role of employer learning in wage dynamics is two-fold. First, the mechanism through which employer learning generates life-cycle wage growth differs from some of the aforementioned theories. Wage growth or decline can occur even if a worker’s productivity does not change. Instead wage changes are generated through changes in employers’ beliefs about a worker’s productivity.⁴ As demonstrated by the theoretical model of [Gibbons and Katz \(1991\)](#), this can be particularly important for understanding wage dynamics associated with flows through unemployment.⁵ If low-productivity workers are more likely to become unemployed and both low and high-productivity workers are indistinguishable at first, then all workers hired from unemployment will receive low wages until their types

¹Dynamic contracting offers an alternative view in which increasing tenure profiles address issues such as principal-agent problems [Lazear and Rosen \(1981\)](#), provide insurance [Harris and Holstrom \(1982\)](#), or screen against adverse selection [Nickell \(1976\)](#).

²Canonical works in human capital theory include [Ben-Porath \(1967\)](#), [Becker \(1975\)](#), [Heckman \(1976\)](#), and [Mincer \(1974\)](#).

³[Jovanovic \(1979\)](#) is the classic reference for employer matching which has recently been extended to job or occupation matching based on skills ([Guvenen et al. \(2015\)](#) and [Lise and Postel-Vinay \(2015\)](#) among others.) Competition generates wage changes in the job ladder model of [Burdett and Mortensen \(1998\)](#) and the sequential auction model of [Postel-Vinay and Robin \(2002\)](#).

⁴This paper will show how learning about fixed productivity contributes to transitory wage dynamics and is not absorbed by individual fixed effects.

⁵[Gibbons and Katz \(1991\)](#) consider only the first separation, whereas I study how employer learning provides repeat separations for select workers.

are revealed to new employers.⁶ Second, the microeconomic literature has provided ample evidence in favor of the employer learning theory.⁷ This paper complements these literatures by quantifying how employer learning at the micro-level contributes to the empirical wage process, with a unique focus on dynamics following an unemployment spell. Additional empirical evidence supporting the importance of employer learning is provided by testing empirical predictions of a structural model.

I begin by building a structural model of employer learning suitable for quantitative analysis. It features workers who differ in their permanent innate productivity which is initially unobservable. Current employers learn their worker’s productivity over the course of the match. If a worker is revealed to be low-productivity, the current employer will select to terminate the match. All workers also face a risk of an exogenous, non-selective separation; for example if an idiosyncratic shock causes their firm shut down. Outside potential employers learn more slowly than current employers. Therefore, they do not know the reason a worker became unemployed: whether she was selectively terminated or separated exogenously at no fault of her own. As a result, unemployed workers of all types are hired at a low wage reflecting the lower average productivity of unemployed workers.⁸ Overtime, new employers learn the productivity of workers they hired from unemployment. Those found to be highly productive will eventually experience wage gains. Those found to be low productivity will be selectively terminated to unemployment yet again. Thus, the theory of employer learning has sharp predictions for subsequent wage and employment dynamics of workers separated to unemployment. It includes the hallmark of imperfect information: while these two paths diverge over time, all workers hired from unemployment are initially indistinguishable from one another.

I provide evidence in support of this theory by testing its predictions related to wage

⁶The implications of this theory for competition and monopsony rents through direct job-to-job changes will be similar to the sequential auction model of [Postel-Vinay and Robin \(2002\)](#) where employer changes precipitate wage changes beyond what can be inferred as changes in productivity.

⁷Evidence comes both from empirical tests as in [Pinkston \(2009\)](#) and [Kahn \(2013\)](#), among others; and also recently through experiments as in [Pallais \(2014\)](#).

⁸This is similar in spirit to [Lockwood \(1991\)](#) who shows screening leads employers to avoid long-term unemployed because selection in re-hiring (versus firing in this paper) leads the long-term pool to be composed of less productive workers.

and employment dynamics of workers separated to unemployment using the Panel Study of Income Dynamics (PSID). I consider unemployed workers who report they were individually separated to be more likely to have been selectively terminated than those who report they were separated by “Firm Shut Down or Closure”.⁹ Consistent with the theory, I find the initial wage losses of individually separated workers are statistically indistinguishable from workers separated by firm closure. Both groups experience wage losses of approximately 17% upon rehire when compared to similar continuously employed workers. Yet, workers separated by firm closure have insignificant wage losses by year 4, as opposed to individually terminated workers who still earn significantly lower wages 15 years later. The data also support predictions for repeat unemployment spells. A worker who is individual terminated has more than double the likelihood of experiencing an individual termination in the following year whereas an observationally equivalent worker separated by firm closure has no higher risk.¹⁰ This further supports the presence of selective termination based on a fixed trait above which can be explained a causal impact of experiencing unemployment on future employment stability.

Having established support for the structural model, I use it to answer the main question: how does employer learning contribute to wage and employment dynamics? I calibrate the model to match separation patterns observed in the PSID for a range of learning speeds supported by the microeconomic literature. In the median calibration, employer learning on it’s own accounts for 78% of the average wage losses following unemployment and replicates divergent paths of individually separated (endogenously selected in the model) and those separated by firm shut down (exogenously separated in the model). It also generates 24% of life cycle wage growth and 13% of wage dispersion observed in the data.¹¹ I provide comparative statics to understand how the speeds of learning contribute to these results. For example, faster current employer learning generally increases cross-sectional wage dispersion, but has ambiguous effects on wage losses associated with unemployment.

⁹This strategy follows [Gibbons and Katz \(1991\)](#).

¹⁰All of these statements are based on regression analysis that control for differences in observable characteristics across the two groups such as industry, education, urban status, etc. (Section 3).

¹¹Robustness providing a range of values for different learning speeds is provided.

These findings indicate that employer learning is an empirically relevant component of wage and employment dynamics, particularly so for workers separated to unemployment. This is significant because this theory provides an unique interpretation of the permanent “individual fixed effects” and transitory components of estimated empirical wage processes. While each worker’s productivity type is fixed, information available about an individual worker’s type is not. It is through this learning process that fixed productivity types have dynamic impacts on workers’ wages over the life-cycle that are not controlled for by individual fixed effects in an empirical regression.¹² The idea that unobservable fixed factors contribute to transitory wage dynamics is meaningful for an array of applications in Labor and Macroeconomics that must take a stand on what drives wage dynamics. For example, the degree of risk implied by wage scars following unemployment is lessened if unproductive workers anticipate they will likely be selectively separated. Furthermore, productivity loss associated with unemployment is as little as zero according to the information theory. This motivates researchers to revisit key questions related to unemployment flows under alternative calibrations of individuals’ risks and productivity dynamics associated with unemployment.¹³

1 Related Literature

The current paper pursues a goal that is distinct, but complementary to the empirical literature conducting statistical tests of whether employer learning generates wage changes.¹⁴ It uses a structural approach to derive quantitative implications of employer learning for over-all life-cycle wage growth and cross-sectional dispersion.¹⁵ The model extends prior theoretic-

¹²I provide regressions on model generated data that include individual fixed effects to make this point.

¹³Macro applications include everything from studying the welfare costs of business cycles (Krebs (2007)) to the persistent effects of aggregate employment fluctuations on aggregate output (Ljungqvist and Sargent (1998)). Particularly relevant policy questions include evaluation of retraining programs, firing costs, efficient unemployment benefits, and more. The literatures on these questions is enormous.

¹⁴Altonji and Pierret (2001) find empirical support for employer learning by comparing specific aptitude test scores not observed by employers to wage patterns. Subsequent literature has attempted to measure learning speed Lange (2007) and test for asymmetry in current and outside employers (Kahn (2013) finds it is asymmetric, Schonberg (2007) finds it is not.

¹⁵This contrast is explained in Pinkston (2009). He provides statistical tests of asymmetric employer learning in an auction framework including strategic behavior closest to the theory in this paper, but concludes that

cal work to provide a multi-period model suitable for quantitative analysis.¹⁶ It develops a simplified learning process and market structure that generates a tractable Bayesian Nash equilibrium and maintains flexibility in learning speed of both current and outside employers over an infinite horizon model.¹⁷

The mechanism through which employer learning generates employment and wage dynamics is similar in spirit to a subset of structural models with completely different micro-foundations. As in the job ladder model of [Burdett and Mortensen \(1998\)](#) and the sequential auction model of [Postel-Vinay and Robin \(2002\)](#), monopsony power allows current employers to pay workers less than their (expected) marginal products until certain frictions are overcome allowing the worker to increase their wage through outside offers. In [Burdett and Mortensen \(1998\)](#) and [Postel-Vinay and Robin \(2002\)](#) monopsony power is delivered by search frictions: there are outside employers willing to pay the worker more, but the worker and these employers cannot find one another. In this paper, monopsony power is delivered by information frictions. Outside employers can find workers, but imperfect information can prevent a job change resulting in wages that depart from both expected and actual marginal product.¹⁸

Finally, this paper relates to the emerging structural literature quantitatively evaluating theories of empirically documented “wage scars” following separation. In this paper, [Jarosch \(2014\)](#), and [Burdett et al. \(2015\)](#) monopsony rents paid to the firm decrease during a continuous employment spell through competition from outside firms, resulting in wage growth for the worker.¹⁹ A restart of this process following unemployment provides initial wage scars, but each theory requires additional ingredients to make scars persistent. In [Jarosch \(2014\)](#) this is provided by job heterogeneity: separated workers select new jobs with high exogenous sepa-

he “does not study the model’s implications for the wage distribution and how it evolves over time or for the probability of turnover, how it evolves, and what it implies for wage changes.”

¹⁶The theory in [Gibbons and Katz \(1991\)](#) that I build on is a three period model. Therefore, they do not examine speeds of wage growth and recovery following unemployment as in this paper.

¹⁷This provides quantitative flexibility in the transmission of employer learning to job mobility and wage growth while allowing employers to behave strategically, the crux of the micro theory literature.

¹⁸I will discuss how superior information of current employers provides monopsony power through a “lemons” market failure of the type studied in [Akerlof \(1970\)](#).

¹⁹This contrasts with theories of lost productivity growth accompanying unemployment such as in [Jung and Kuhn \(2016\)](#), [Ljungqvist and Sargent \(1998\)](#), and [Krolikowski \(2017\)](#).

ration rates. In the current paper, long-run divergence is provided by worker heterogeneity which is initially unobservable. [Burdett et al. \(2015\)](#) use both worker and job heterogeneity. They also implement a job-ladder model but consider a calibration in which workers are of two fixed types: low and high skilled, to generate dispersion in post-unemployment outcomes.

2 Model of Asymmetric Employer Learning

I construct a quantitative model of labor markets where wage and employer changes are driven by strategic interactions of workers and employers under asymmetric learning. I begin with a static version of the game. I use it to intuit the equilibrium strategies of players and the qualitative predictions they deliver. I then present the infinite horizon model used for quantitative analysis.

2.1 Static Model

The world is populated by a unit measure of workers who are either currently matched or unmatched; and a large mass each of two-types of employers: incumbents and poachers. Incumbents are the employers currently matched with a worker. Poachers are outside potential employers looking to steal a currently matched worker. Unmatched workers are offered wages through Bertrand competition of many employers.²⁰

Players *Workers.* Workers are characterized by fixed ability $i \in \{\ell, h\}$, low or high, drawn independently across workers: $i = h$ with probability π ; and $i = \ell$ with probability $1 - \pi$. *Employers.* Employers operate identical production technologies each employing a single worker. High ability workers produce an output of one with any employer ($y_h = 1$) and low ability workers produce output of zero ($y_\ell = 0$).

Actions. Each period, incumbent and poaching employers submit a wage offer to currently

²⁰Bertrand competition of identical outside employers is a modelling tool to restrict the set of equilibria used in the asymmetric learning literature in Personnel Economics (see: [Lazear and Oyer \(2007\)](#) and specifically, [Waldman \(1984\)](#).) I omit formal proofs that Bertrand competition between identical poachers delivers wages that satisfy a zero expected profits condition because this is well known.

matched workers. Denote poaching employers with a star. Wage offers are subject to a wage floor (\underline{w}) higher than each the flow utility of nonemployment (zero) and the output of a low-ability worker.²¹

- *Incumbent action*: offer a wage $w \in [\underline{w}, \bar{w}]$
- *Poachers action*: offer a wage $w^* \in [\underline{w}, \bar{w}]$.
- *Worker action*: Quits to a poacher ($Q = 1$) or stays with incumbent ($Q = 0$)

Payoffs.

- *If the worker does not quit*: the worker receives payoff w ; and the incumbent receives payoff $y_i - w$.
- *If the worker quits*: the worker receives payoff w^* ; and a randomly assigned poacher receives payoff $y_i - w^*$.

Information. A worker's ability is her private information. The incumbent begins the game with a rational prior $p = \pi$: the probability a random worker is the high ability type. At the beginning of the period, a measure d of workers are randomly separated to unemployment (the match “shuts down”). Next, before wage offers are made, the incumbent learns their worker's ability perfectly with probability μ . If the incumbent learns their worker is high ability, they update their prior to $p = 1$. If they learn their worker is of low-ability, the match is unproductive and the worker is separated to unemployment. A poacher's rational prior about a worker is denoted as p^* . With probability ν each poacher is informed of *the incumbent's information* and adopt the same prior as the incumbent.²² Otherwise, they are “uninformed” and only see the worker is employed or unemployed. Uninformed poachers rationally adopt

²¹In the dynamic model, the wage floor will prevent employers from offering contracts only the high type would accept. If such separating contracts exist (it depends on the parameterization) they would involve a wage starting below the flow value of the outside option of employment and increase over tenure. Since low-type workers are more likely to be fired before they achieve wage increases with tenure, they would value the job less than high-type workers and could be separated. This is related to the single-crossing property in many applications of separating contracts with adverse selection.

²²Crucially, poachers never have information about the worker that the incumbent does not.

prior $p^* = p^e = \frac{(1-d)\pi}{(1-d)(1-\mu(1-\pi))}$, the probability a random employed worker is the high type. Poachers competing for an unemployed worker have rational prior $p^* = p^u = \frac{d\pi}{d+\mu(1-\pi)}$, the proportion of unemployed that are the high type. Critically, all poaching employers have the same information about a worker and the incumbent does not observe the action of the poachers or whether they are informed. The worker observes all actions, information, and states.

Beliefs and Strategies. Strategy profiles are a mapping from a player's state to their action, taking the strategies of other players as given.²³ I consider pure strategies defined as:²⁴

$$\begin{array}{ll} \text{Incumbent} & w(p) \in [\underline{w}, \bar{w}] \\ \text{Poachers} & w^*(p^*) \in [\underline{w}, \bar{w}] \\ \text{Worker} & Q_i(p, p^*) \in \{0, 1\} \end{array}$$

An incumbent rationally believes the poachers will play the informed wage offer with probability ν and the uninformed offer with probability $1 - \nu$. Informed poachers know the wage the incumbent will offer since we only consider pure strategies. Lastly, uninformed poachers have rational beliefs over the distribution of worker-incumbent matches they may meet according to the equilibrium distribution $\{\lambda^e(p)\}$ of incumbent priors $p \in \{\pi, 1\}$.²⁵

A worker's decision to accept a wage offer could potentially reveal information about their true type. As such, the employer whose offer the worker accepts updates their belief about the worker's type based on their prior, the wage the offer, and the fact the worker accepted it. Denote the Bayesian posterior as $B(p, w)$ for the incumbent if the worker stays and $B^*(p^*, w^*)$ for the poaching employer the worker joins if she quits.

Timing.

Workers begin life as new entrants matched with an incumbent employer. Nature ran-

²³I drop typical notation including the strategy profile of other players as a notational convenience.

²⁴Bertrand competition by the poachers yields a unique wage offer given their common information. This results in a one-to-one mapping from p^* to w^* . The incumbent plays a pure strategy and so there is also a one-to-one mapping from p to w . Together, this allows the state of the worker to be summarized by the priors (p, p^*) , omitting w and w^* as redundant.

²⁵Since incumbents are rational, the true distribution of types is correctly summarized by the prior.

domly separates a share d of the workers to unemployment. If a worker is not separated, her incumbent learns her type with probability μ . Workers revealed to be low type are chosen to be separated to unemployment by their employer. Poachers become informed of the type of a worker remaining employed with probability ν . Next, incumbents and poachers simultaneously submit wage offers to employed workers. Workers choose whether to take a poachers' offer and quit or stay with the incumbent. Unemployed workers accept a wage offer from many Bertrand competing employers. Production takes place, players collect payoffs, and the game ends.

Definition 2.1 (Symmetric Perfect Bayesian Equilibrium- Static Model). An equilibrium Φ is a set of pure strategies: $w(p) \in [\underline{w}, \bar{w}]$ wage offer of incumbent; $w^*(p^*) \in [\underline{w}, \bar{w}]$ wage offer of poachers; $Q_i(p, p^*) \in \{0, 1\}$ quit strategies of workers²⁶; posterior belief functions $B(p, w') \in [0, 1]$ and $B^*(p^*, w^*)$; measures $\{\lambda^e(p)\}$ of workers with incumbent prior p ; and λ^u proportion of unemployed that are high type such that, given exogenous parameters Θ :

- Strategies are sequentially rational given Φ
- Priors are consistent with equilibrium distributions
- Posteriors are Bayesian where possible.²⁷
- Distributions $\{\lambda_i^e(p)\}$ and λ^u are consistent with parameters and optimal strategies.
- The worker stays with the incumbent when she is indifferent.²⁸

Equilibrium Characterization- Static Game. The equilibrium, to be verified, is characterized as:

Incumbent $w(p) = \min\{p, \min\{w^*(p), w^*(p^e)\}\}$; the incumbent wage equals the minimum of the informed and uninformed poacher's offers up to the worker's expected output.

²⁶The definition of a symmetric equilibrium is that all agents within one *type* play the same strategy.

²⁷I define off-path beliefs to be consistent with the limit of the workers' forced mixed strategies: $\lim_{\epsilon \rightarrow 0} Q_{i\epsilon}$ where $Q_{i\epsilon} \in [\epsilon, 1 - \epsilon]$. However, this equilibrium can be supported by alternative equilibrium refinements including Cho-Kreps Intuitive Criterion, Sequential Equilibrium as Kreps and Wilson, and is the limit of Control Cost equilibria as Selton.

²⁸To prevent equivalent wage paths dynamics with superfluous turnover.

Poachers • Informed: $w^*(p) = p$ the informed poachers pay a wage equal to the worker's expected output.

- Uninformed: $w^*(p^e) = \min_{p \in \mathcal{P}} \{w(p)\}$ uninformed poacher pay the lowest incumbent wage offered in the economy.

Workers For all $i \in \{\ell, h\}$: $Q_i(p, p^*) = 1$ iff $w^*(p^*) > w(p)$. Workers of each type play the same strategy (pooling): quit if the poachers' wage is higher than the incumbent's.

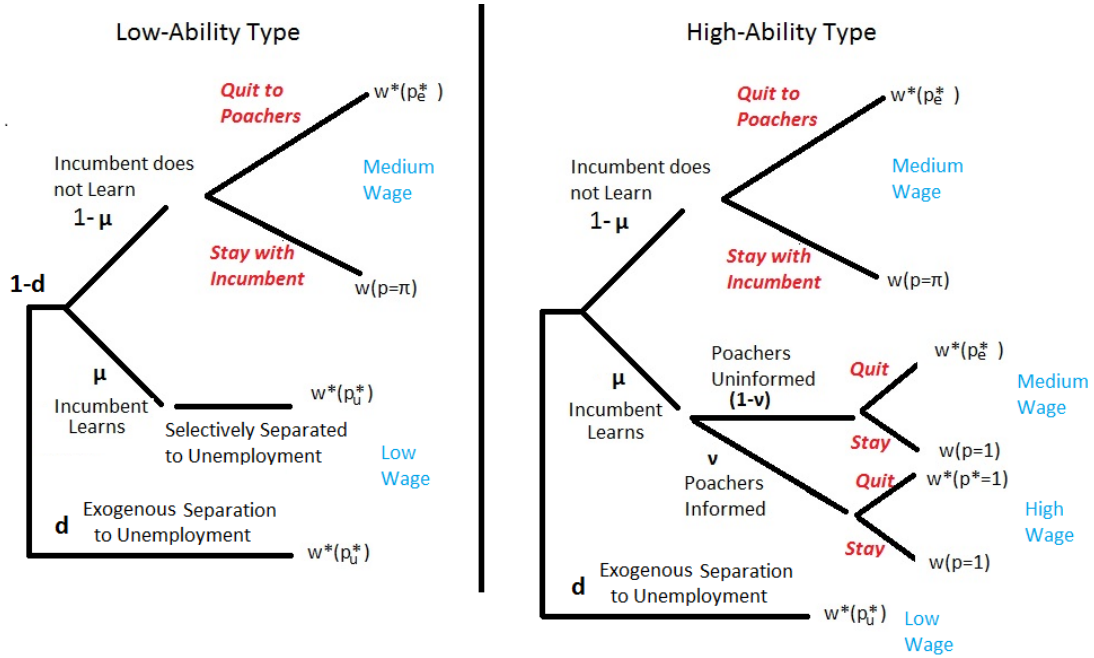


Figure 1: Possible outcomes for low and high-ability types. Workers have full information of which node they are at and their choices are in red. Equilibrium wage outcomes (verified below) are shown in blue. If the incumbent has imperfect information (top branch) each type is treated identically, but if the incumbent learns (middle branch) outcomes of the two types differ. Inferior information of outside employers provides identical outcomes to exogenously separated high-types as those low-types selectively separated when their past employer learned. However, in the repeated game of the next section high type workers can recover to high wages as their next incumbent employer and poachers learn their type whereas this path is not available to low-types who will be selectively separated when their next employer learns their type (Figure 3).

I now verify these strategies are an equilibrium. The strategy of the workers is obvious. Taking others' strategies as given, they maximize their pay-off by choosing the highest offered

wage. The strategy of informed poachers is set by Bertrand competition. They offer a wage that leaves them zero expected profits.²⁹

Incumbents know poachers will play one of two wages: informed $w^*(p)$ with probability ν or uninformed $w^*(p^e)$ with probability $1 - \nu$. Given the worker's strategy to accept the highest wage offer, the incumbent will keep the worker if they match or beat the poachers' offer. Informed poachers share the incumbent's information and Bertrand compete any potential profits to zero, leaving no opportunity for the incumbent to profit. The incumbent only has an opportunity to earn positive profits if poachers are uninformed and offer a wage less than the expected output: $w^*(p^e) < p$. In this case, the incumbent will choose $w(p) = w^*(p^e)$. They offer the worker as little as possible to keep the worker from the uninformed poacher and earn profits $p - w^*(p^e) > 0$ if poachers are uninformed and zero if they are informed.

Proposition 2.2. *The incumbent's strategy is characterized as matching the minimum poacher offer conditional on it being less than the incumbent's expectation of the worker's output:*

$$w(p) = \min\{p, \min\{w^*(p), w^*(p^e)\}\} \quad (2.1)$$

Proof. The incumbent's expectation of the worker's productivity is p . Informed poachers play $w^*(p) = p$. Then any $w > w^*(p)$ cannot be optimal; the incumbent will earn negative profits. If $w^*(p) > w^*(p^e)$, the incumbent will keep the worker if the poachers are uninformed (probability $1 - \nu$) if they pay any $w \in [w^*(p^e), w^*(p)]$. This yields an expected profit $\max_{w \in [w^*(p^e), w^*(p)]} (1 - \nu)[p - w]$. Clearly, the maximizing solution is $w(p) = w^*(p^e)$ \square

When uninformed poachers Bertrand compete, they offer a wage that yields zero expected profits *conditional* on the type of worker who accepts it. Given the strategy of workers, they will only successfully poach a worker if they offer a wage higher than the incumbent. The superior information of the incumbent leads to a classic “lemons” problem. The incumbent's optimal strategy is to match offers up to the worker's expected output. Therefore, if poachers offer a higher wage than the incumbent, they will get the worker at a price greater than

²⁹Bertrand competition follows from the assumption that there are many identical poachers, all sharing the same information and production technologies.

her expected output. In equilibrium here, as well as in the dynamic game, the uninformed poachers offer the lowest incumbent wage offered in the entire economy and never poach any workers to avoid overpaying; a typical “lemons market failure”.³⁰

Proposition 2.3. *The uninformed poachers’ strategy is to match the lowest incumbent wage offer in the economy, given the equilibrium distribution of incumbent-worker matches $(\{\lambda^e(p)\})$.*

$$w^*(p^e) = \min_{p \in \mathcal{P}} w(p) \quad (2.2)$$

Proof. Incumbents can be of two types: $p \in \{\pi, 1\}$, those who learned the worker is the high type ($p = 1$) and those who have not ($p = \pi$). The theorem states the optimal strategy of uninformed poachers is to offer $w^*(p^e) = \pi$. Suppose not. Consider $1 > w^*(p^e) > \pi$. Given the incumbent’s strategy $w(p) = \min\{w^*(p^e), p\}$, the uninformed poacher will only poach from incumbents who have not yet learned their worker’s type ($p = \pi$). This delivers negative expected profits: $\pi - w^*(p^e) < 0$. Similarly, if $w^*(p^e) > 1 > \pi$, the uninformed poachers poach all workers at negative expected profits: $(\pi(1 - \mu) + \mu - w^*(p^e)) < 1 - w^*(p^e) < 0$. These strategies are dominated by $w^*(p^e) = \pi$, earning zero profits.³¹ \square

Employers Bertrand compete for unemployed workers. They offer a wage equal to the expected output of a worker hired from unemployment: $w(p^u) = \frac{d\pi}{d + \mu(1 - \pi)}$. This expectation is rational: $d\pi$ is the flow of high ability workers to unemployment through exogenous shut downs and the denominator, $d + \mu(1 - \pi)$ is total flows d of all workers to unemployment through shutdowns plus an additional flow $\mu(1 - \pi)$ of low type workers whose employer learns their type and fires them.

To establish the uniqueness of the equilibrium, we finally need to rule out signalling strategies by showing each type of workers play the same strategy conditional on the state p, p^* .

³⁰See the seminal paper [Akerlof \(1970\)](#).

³¹The alternative, $w^* < \pi$ fails the equilibrium refinement for off-path outcomes.

Proposition 2.4. *Workers of each type play the same strategy (pooling).*

$$Q_h(p, p^*) = Q_\ell(p, p^*) \quad \forall p \in \mathcal{P} \times \mathcal{P}^*. \quad (2.3)$$

Proof. Both types meet the uninformed poachers, but if only one type quits the uninformed poachers will perfectly know which type of worker they are getting. Case 1: suppose only low-ability types quit to the uninformed poachers. This is not optimal because it reveals their type, resulting in being fired and earning the lowest wage in the economy ($w^*(p^u)$). Case 2: suppose only high-ability types quit to the uninformed poachers. This changes the posterior of the uninformed poachers to know the worker they get is high type with certainty. Since they Bertrand compete, the uninformed poacher wage becomes $w^*(p^e) = 1$. However, at this wage, they will poach low type workers as well. This contradicts the uninformed posterior that only high types accept their offer and $w^*(p^e) = 1$ no longer satisfies zero expected profits, a contradiction. \square

Discussion The equilibrium displays two properties key to tractability in the quantitative model. First, it is a pooling equilibrium: workers of each ability type $i \in \{\ell, h\}$ play the same strategy. Pooling rules out strategic behavior of employers to only attract one type. Second, due to a classic “lemons market failure”, uninformed poachers never successfully poach a worker from an incumbent. Together, these properties imply employers do not receive new information from workers’ actions; ie: their Bayesian posterior equals their prior.

2.2 Dynamic Quantitative Model

The dynamic model is essentially a repetition of the static game as a Markov game.³² Information is carried perfectly from one period to the next, summarized in a Markovian fashion by the incumbent’s belief p and poachers’ belief p^* , each a probability the worker is the high-type.³³

³²See the extended appendix for the full formal definition of the game.

³³One could also imagine additional variables in the information set including age and tenure. Incorporating these features, particularly tenure is non-trivial. Workers may try to manipulate their tenure to signal their type. We are no longer assured of a pooling equilibrium.

The dynamic model crucially maintains a tractable information structure with a finite set of beliefs in equilibrium. This is provided by the following assumptions and results. During each period of employment: (i) the incumbent employer learns the worker's type with probability μ ; and (ii) outside potential poaching employers become informed of the current employers' information with probability ν . The Poisson structure implies the speed of employer learning is independent of the worker's tenure. This assumption will provide pooling over ability in workers' strategies. The assumption that many poachers share weakly less information than incumbents will imply no strategic learning in equilibrium.³⁴ Given these assumptions, the in-consequence of uninformed poachers is maintained and the Bayesian posterior of the employer who wins a worker equals their prior.

The model is modified to be appropriate for quantitative analysis. Flexibility in the productivity of each type is assumed: $y_h > y_\ell = 1$. Second, employers' wage offers are restricted downward ridged wages; ie: an incumbent cannot cut the wage of their employee in subsequent periods. This assumption is chosen purely to prevent superfluous transitions to poaching employers that raise a worker's wage for one period then reduce it.³⁵

The three relevant state variables are: current incumbent and poachers' priors $(p, p^*) \in \mathcal{P} \times \mathcal{P}^*$, and the standing wage $w \in [\underline{w}, \bar{w}]$ (if the worker is employed).

Workers have time-separable linear utility over wages and employers maximize expected future profits. Workers and employers share discount factor $\rho = 1 - d - \delta$. Payoffs are:

$$\begin{array}{ll} \text{Workers} & \mathbf{E}_i \left[\sum_{t=0}^{\infty} \rho^t w_{it} \right] \\ \text{Employers} & \mathbf{E} \sum_{t=0}^{\infty} \rho^t [y_i - w_{it}] \end{array}$$

Workers begin life as new entrants with their ability as their private information. Employers simultaneously offer new entrants wage contracts. During their career, the static game

³⁴As opposed to a world where poachers learn through independent signals that may contain different information than the current employer.

³⁵Firings do not occur because of downward rigid wages. They occur because it is assumed there is a minimum wage sufficiently high such that that low ability unskilled workers deliver negative expected profits.

repeats each period the worker is employed, unless the worker dies with probability δ in which case she is replaced next period by a new entrant. Wages are paid and profits collected at the end of each period. The explicit timing is:

1. *Employer Learning*

- *Incumbent Learning*: probability μ incumbent learns the worker's type.
- *Poacher Learning*: probability ν poachers informed of incumbent's information.
- *Selective Separation*: Low-ability workers are fired if incumbent learned type.

2. *Wage Offers & Quits*

- *Incumbent wage offers*: choose $w'(p, w) \geq w$
- *Poachers' wage offers*: $w^*(p^*)$ determined by Bertrand Competition.
- *Employer to Employer Quits*: Employed workers choose whether to quit to poachers.
- *Bertrand competition for new entrants and unemployed workers*.

3. *Production* Wages paid and profits collected.

4. *Nature*

- *Exogenous Separation (Shut Down)*: Match ends randomly with probability d .
- *Worker Death, Birth*: Measure δ die, replaced by new entrants

Definition 2.5 (Symmetric Perfect Bayesian Equilibrium- Dynamic Model). An equilibrium Φ is a set of pure strategies: $w'(p, w) \in [w, \bar{w}]$ wage offer of incumbent; $w^*(p^*) \in [\underline{w}, \bar{w}]$ wage offer of poachers; $Q_i(p, w, p^*) \in \{0, 1\}$ quit strategy of workers; posterior belief functions $B(p, w) \in [0, 1]$ and $B^*(p^*, w^*) \in [0, 1]$ each defined over finite sets $p \in \mathcal{P}$, $p^* \in \mathcal{P}^*$, and their cross products as appropriate; stationary measures $\{\lambda_i^e(p, w)\}$ of employed workers of type i with incumbent prior p and wage w with support Ψ ; and stationary distribution $\{\lambda_i^u\}$ of workers of type i flowing through unemployment each period; such that, given exogenous parameters Θ :

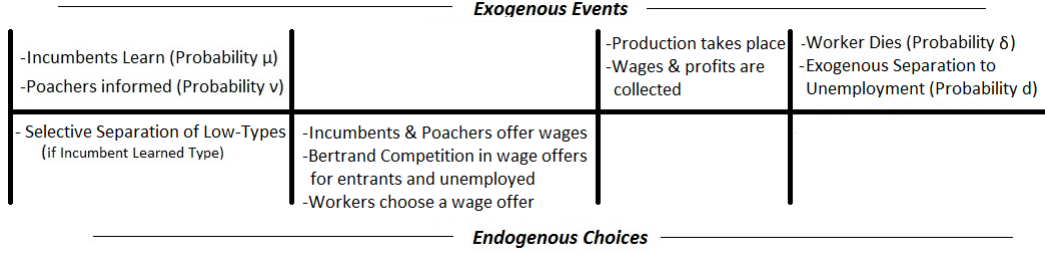


Figure 2: Sequence of events in one period of the dynamic game. Endogenous choices listed below timeline. Exogenous events listed above timeline.

- Strategies are sequentially rational given Φ
- Priors are consistent with equilibrium distributions $\lambda_i^e(p, w), \lambda_i^u$.
- Posteriors are Bayesian where possible.³⁶
- Workers stay with the incumbent when indifferent.
- Stationary distributions $\{\lambda_i^e(p, w), \lambda_i^u(p)\}_{i \in \{\ell, h\}}$ are consistent with strategies; and the measure of unemployed are $u = \lambda_\ell^u + \lambda_h^u$.

Equilibrium Characterization- Dynamic Game Denote $\Omega(p, w)$ as the value to an employer with a worker with an interior prior $p \in (0, 1)$ (ie: has not yet learned the worker's type), and standing wage w ; and $\bar{\Omega}(p, w)$ as the value to an employer who knows their worker type with certainty (ie: $p \in \{0, 1\}$). Also denote p^u as the proportion of workers in unemployment that are the high type. The equilibrium is characterized by the following strategies:

- *Incumbent.* $w'(p, w) = \min\{y(p), \max\{w, w^*(p^e)\}\}$; the incumbent pays the worker as little as possible to keep her if poachers are uninformed, conditional on this value not exceeding the expected output given the incumbent's prior.

³⁶I define off-equilibrium path beliefs as the limit of the workers' forced mixed strategies. However, this equilibrium can be supported by many off-equilibrium path beliefs and satisfies other equilibrium refinements including Cho-Kreps Intuitive Criterion, Sequential Equilibrium as Kreps and Wilson, and is the limit of Control Cost equilibria as Selton.

- *Informed Poachers.* $w^*(p)$ is such that $\Omega(p, w^*(p)) = 0$ if $p \in (0, 1)$ and $\bar{\Omega}(p, w^*(p)) = 0$ if $p \in \{0, 1\}$; informed poachers' offer satisfies zero expected profits.
- *Uninformed Poachers.* $w^*(p^e) = \min_{p \in \mathcal{P}} \{w(p)\}$; uninformed poachers offer the minimum incumbent wage offer to employed workers.
- *Workers.* $Q_i(p, w, p^*) = 1$ iff $w^*(p^*) > w'(p, w)$ Quit to poachers if their offered wage is higher than the incumbent.

The crucial component is that the pooling equilibrium in workers' strategies is maintained. The intuition for this result is that it is costless for low types to mimic high types. This is because the only type-dependent payoff for the low type occurs when the incumbent learns the worker's type. Poisson learning provides that this payoff is independent of her past actions, particularly her current tenure with the incumbent. This can be seen formally in the workers' value functions, taken other strategies as given. Denote $\bar{V}_i(w, p^*)$ as the value for a worker whose incumbent has learned her type and $V_i(p, w, p^*)$ as the value for a worker whose incumbent has not learned her type.

$$\bar{V}_i(w, p^*) = \begin{cases} (\rho - d) \frac{w + \rho \nu \frac{y_h}{1-\rho}}{1-\rho(1-\nu)} + d \frac{\theta_i V_i(p^u, w^*(p^u), p^*)}{1-\rho(1-\theta_i)} & \text{if } i = h \\ 0 + \frac{\theta_i V_i(p^u, w^*(p^u), p^*)}{1-\rho(1-\theta_i)} & \text{if } i = \ell, \text{ worker will be fired} \end{cases}$$

Where the first term is the present value of the current match (zero for low-types who will be selected for separation) and the second term is the present value of unemployment.

$$V_i(p, w, p^*) = \begin{cases} (\rho - d) \frac{w + \rho \mu \bar{V}_h(w'(1, w), p^*)}{1-\rho(1-\mu)} + d \frac{\theta_i V_i(p^u, w^*(p^u), p^*)}{1-\rho(1-\theta_i)} & \text{if } i = h \\ (\rho - d) \frac{w + \rho \mu \bar{V}_\ell(w'(0, w), p^*)}{1-\rho(1-\mu)} + d \frac{\theta_i V_i(p^u, w^*(p^u), p^*)}{1-\rho(1-\theta_i)} & \text{if } i = \ell \end{cases}$$

Observe incumbent belief p has no differential impact on the value of a match to each type of worker ($V_i(p, w, p^*)$). $V_i(p, w, p^*)$ is clearly increasing in w and so the argument of Proposition 2.4 applies.

Proposition 2.6. *Workers' quit strategies exhibit pooling over types: $Q_h(p, w, p^*) = Q_\ell(p, w, p^*)$ for any $w \in [\underline{w}, \bar{w}]$ and all $(p, p^*) \in \mathcal{P} \times \mathcal{P}^*$.*

Proof. Application of Proposition 2.4. □

Given the worker's strategy to quit only to informed poachers, regardless of type, the remaining value functions are straightforward. Working backwards from the value to an employer who has learned his worker's type:

$$\bar{\Omega}(p, w) = \begin{cases} \frac{y_h - w}{1 - \rho(1 - \nu)} & \text{if } p = 1 \\ 0 & \text{if } i = \ell, \text{ worker will be selectively separated} \end{cases} \quad (2.4)$$

In first case the worker is the high-type and the value is the present discounted profits until the match ends when poachers become informed. In the second case, the incumbent learns the worker is low-ability and selects to end the match that period and earn zero profits. Given $\bar{\Omega}(p, w)$, the values before learning are:

$$\Omega(p, w) = \frac{y(p) - w + \rho\mu[p\bar{\Omega}(1, w'(1, w))]}{1 - \rho(1 - \mu)} \quad (2.5)$$

Joint-worker turnover and wage paths are shown in Figure 3. The mechanisms in the static model repeat causing the high types to recover over time and low-types will be repeatedly separated. All else is the same: unskilled, low-ability workers still are fired to unemployment when the incumbent learns; exogenous shut-down provides additional flows to unemployment of all types of workers; all workers hired from unemployment earn identical wages equal to the lowest in the economy, direct employer to employer changes occur jointly with a wage increase when outside poachers become informed of the incumbent's information.

The Cases of Purely Private or Symmetric Public Learning. If learning is purely private ($\nu = 0$), current employers always have an informational advantage over outside, potential employers. This shuts down poaching ("lemons market failure") and gives current employers full monopsony power. Information rents exist and will be paid to the incumbent in

Typical Paths of Individually Selectively Separated and Exogenously Separated Workers

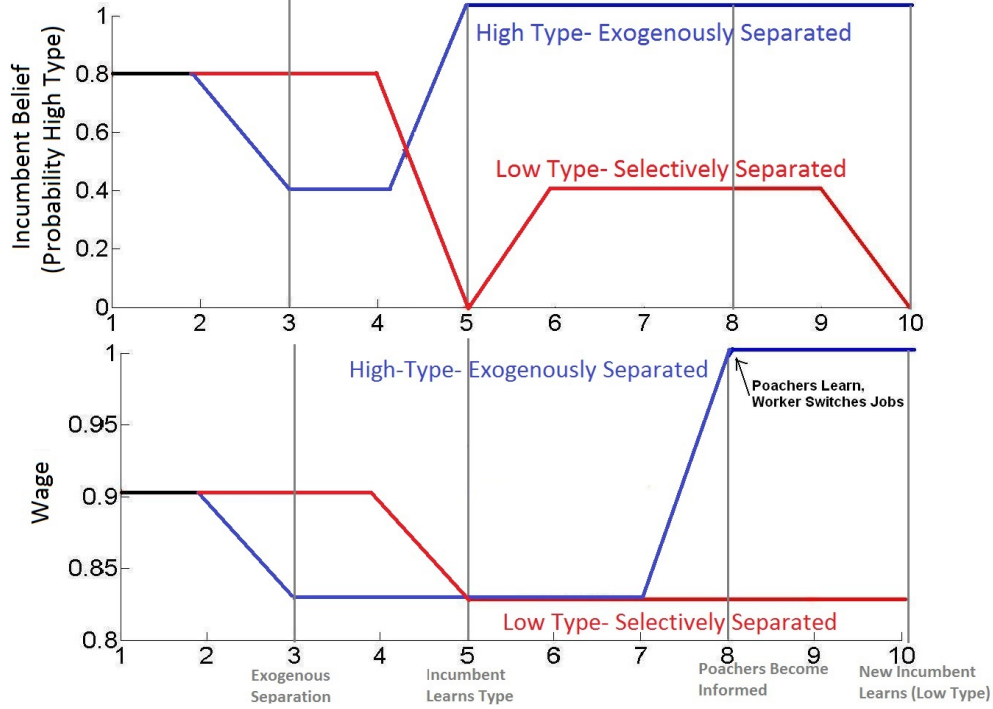


Figure 3: Example wage and belief paths of a low and high ability new entrant who experience each type of separation. The share of high-types in the economy and in unemployment are 0.8 and 0.4, respectively. Wage levels are not from a specific calibration. Observe low-types never recover after they are separated (for whatever reason- selective or not): low wages persist accompanied by repeat future selective separations. On the contrary, high-types who are exogenously separated maintain stable employment and eventually recover to high wages as their subsequent employers learn their type

their entirety. Alternatively, if learning is public ($\nu = 1$), there are no information rents.³⁷ The current employer will immediately lose the worker at the standing wage.³⁸ The implications of these results are as follows. Public learning implies no workers revealed to be low-type will ever be hired again. This means the pool of unemployed will actually be *better* than the population of new entrants through attrition of low-types throughout the life course. We would

³⁷One can see these results in Equation 2.4. At $\nu = 0$ the continuation value at the standing wage is the present discounted value until the match ends exogenously (ρ): $\frac{y_h - w}{1 - \rho} > y_h - w$. At $\nu = 1$, the continuation value at the standing wage is $y_h - w$; the worker will be poached before the next period.

³⁸Or they can increase the wage to equal her expected output y_h to keep her and earn the same continuation value: zero.

expect to see wage *gains* following unemployment. On the other hand, purely private learning $\nu = 0$ delivers no wage increases for employed workers. This means that *all* workers hired from unemployment will be paid $w(p^u)$ for the rest of their career. There will be no recovery for exogenously separated workers. They will be identical to the endogenously selectively separated.

3 Empirical Evidence

The model provides sharp predictions for the paths of workers selectively separated to unemployment and those separated exogenously at no fault of their own. In this section, I analyze PSID data to test the model’s predictions for both wages and the likelihood of subsequent separations following each type of initial separation.³⁹ First, I map each type of separation in the model to the PSID data following [Gibbons and Katz \(1991\)](#). I consider only separations that the respondent reports as “involuntary”.⁴⁰ I further restrict this group to individuals who report that their prior job ended by one of the following categorized reasons: (1) “Company folded / changed hands / moved out of town; employer died / went out of business”; (2) “Laid off; fired”.⁴¹ The former group is labelled as “shut down” workers and the later group is labelled as “individually separated” workers. The former/latter are empirical counterparts to the exogenous/selectively separated in the model.⁴²

I begin by testing the model’s implications for wage dynamics: (i) the impact of a separation on wages should initially be the same regardless of reason for separation; (ii) individually separated workers should never recover to the reference group of never separated workers, in

³⁹Full details on sample statistics, construction of variables, and full regression results are available in the extended appendix.

⁴⁰An involuntary separation is identified as follows. First, the individual must either be unemployed (report they are looking for work) or have less than one year tenure on their current job. This is double checked by verifying the individual has less than 52 weeks of employment in the reference year. Second, the individual must report that their prior job ended by one of the following categorized reasons: (1) “Company folded / changed hands / moved out of town; employer died / went out of business”; (2) “Laid off; fired”.

⁴¹These categories represent less than a third of the responses to “what happened to your prior job”. Other answers span “strike, lockout”, “quit”, “first full time job”, “was self employed before”.

⁴²They need not all be truly selectively separated, but they must have a greater share of workers selectively separated than the second group of “Shut down” workers.

contrast to those separated by shut down who should; (iii) long-run outcomes of individually separated workers should differ from those separated by shut-down on account of selection in the former group. I use the event study framework of [Jacobson et al. \(1993\)](#) to estimate the effect of a separation at time $t - n$ on current wages w_t of individual i according to this regression:

$$\ln(w_{it}) = \gamma X_{it} + \gamma_1 \exp_{it} + \gamma_2 \exp_{it}^2 + \sum_{n=-1}^{14} \beta_n D_{nit} + \delta_t + \alpha_i + \epsilon_{it}$$

The dependent variable is the natural logarithm of hourly earnings. The independent variables included in the PSID regression that are also included in the model data regression are: a quadratic of experience, D_{it}^n a dummy variable that captures a separation in year $t - n$, and individual fixed effects α_i .⁴³ The separation dummy variable D_{it}^n includes dummies for one year prior to, the year of, and each of the first through fifteenth years following the separation (ie: $n \in \{-1, 0, 1, 2, \dots, 15\}$).⁴⁴ Additional variables include time fixed effects δ_t , a quadratic in experience \exp_{it} , and additional controls X_{it} following the employer learning literature (specifically [Pinkston \(2009\)](#)).⁴⁵

The estimated effects of a separation on wages, stratified by reason for separation, are shown in Figure 4.⁴⁶ Wald tests on the coefficients verify the model's predictions. First, the typical separated worker receives a wage loss on her next job and the magnitude of this loss is not impacted by the type of separation. The test of the null hypothesis that regression coefficients for each type of separation are identical in period $t + 1$ has a p-value=0.6470. Second, the duration of these losses vary by reason for unemployment. The test of the null hypothesis that the coefficients on each type of separation is the same first fails at the 5%

⁴³Wage growth from experience is provided by employer learning in the model. Individual fixed effects are provided by heterogeneous ability types.

⁴⁴While firm fixed effects are not available in the PSID, the leading dummy (one-period before a separation) captures some of the impact of heterogeneity in firms likely to close.

⁴⁵They include dummies for each: old age (greater than 55 years old), blue collar job occupation, union status, manufacturing industry, marital status, highest education is a high school degree, highest education is a college degree, residence in city of population greater than 100,000, residence in city of population less than 25,000.

⁴⁶The plots depict the coefficients on the separation dummies from this regression. I list the values of the regression coefficients from this specification and from a specification where both groups: shut down and individually separated are pooled into all involuntary separators.

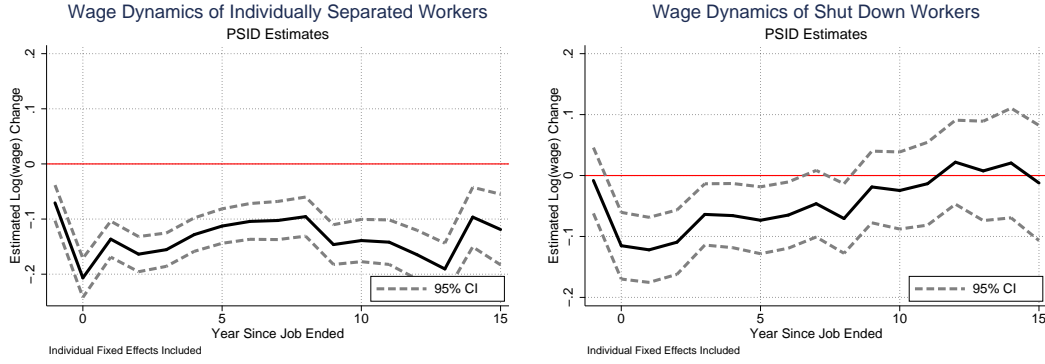


Figure 4: Wage Dynamics of Separated Workers

level in year 3. Wages of shut down workers recover to the control group of continuously employed workers in the 7th year following their unemployment spell. The test of the null hypothesis that the coefficient equals zero cannot be rejected the 5% level (p-value=0.0975). Wages of individually separated workers never recover. The null hypothesis that the coefficient after an individual separation is zero even fourteen years later is rejected at the 1% level (p-value=0.005).

Next, I test the model's predictions for repeat separation patterns stratified by reason for initial separation.⁴⁷ The three key predictions for the impact of a separation in year $t-1$ on the likelihood of a separation in t are: (i) an individual selective separation in $t-1$ should increase the probability of separation in t ; (ii) a non-selective exogenous separation in $t-1$ should have no impact on the likelihood of selective separation in t ; (iii) neither type of separation should affect the likelihood of a non-selective exogenous separation (shut down) in t .

I estimate the impact of a separation on the likelihood for a given type of separation in the following year according to a logit specification:

$$D_t^r = \gamma X_{it} + \gamma_1 age_{it} + \gamma_2 age_{it}^2 + \beta_{-1}^{SD} D_{t-1,i}^{SD} + \beta_{-1}^F D_{t-1,i}^F + \delta_t + \alpha_i + \epsilon_{it}$$

⁴⁷Stevens (1997) first showed repeat separations accompanied the wage losses of separated workers. I build on her work by documenting differences between shut-down and individually fired workers and providing a theoretical model to account for these differences.

The independent variables are identical to the wage regression with the exception that only a one period lag in the indicator variable for an individual separation ($D_{t-1,i}^F$) or one by firm closure ($D_{t-1,i}^{SD}$) are used, and there are no fixed effects. The two outcome variables are dummies for the incidence of an individual separation D_t^F or a separation due to shut down D_t^{SD} in year t . Table 1 presents the marginal impact of each type of separation in period $t - 1$ on these outcomes. The impact of an individual separation in $t - 1$ has a large and significant impact on the likelihood of an individual separation in t (the hypothesis of no effect is rejected at 1%), but an exogenous shut down is precisely estimated to be zero. Further, the odds ratio of this likelihood is 1.92 (s.e. 0.1160; p-value 0.000) if an individual separation was experienced in $t - 1$ versus a shut down. The evidence is less strong for the third prediction. We fail to reject that either type of separation affects the likelihood of a separation due to shut down in the following year.⁴⁸ These results bolster the claim that workers separated through these two channels differ permanently in a fundamental way beyond that which can be attributed to a causal effect of unemployment on future separations.

Comparison to Prior Empirical Studies. The sample selection follows the employer learning literature. It contrasts with the sample selection common in the worker “displacement” literature.⁴⁹ For example, [Jacobson et al. \(1993\)](#) and follow ups including [Couch and Placzek \(2010\)](#) restrict their sample to prime age high-tenured male workers with high labor-force attachment. The definition of “displacement” also differs from “separation”. It is restricted to large employers (greater than 50) that terminated more than 30% of their workforce at the establishment. With such restrictions they are able to find persistent losses for this select group of displaced workers. However, within the “displacement” literature, [Flaen et al. \(2013\)](#) show workers separated from distressed firms recover more quickly than those separated for other reasons. This is a similar result to the “individual separation” versus

⁴⁸This is a curious finding that the theory in this paper is silent on. However, we do fail to reject that the impact of each separation is the same at the 5% level as the theory suggest, although the same cannot be said at the 10% level.

⁴⁹The sample chosen is appropriate for the question at hand. It is more inclusive and relevant for understanding aggregate worker flows and wage dispersion because these individuals account for a larger share of the workforce. The inclusion of young workers at the start of their career is also important for disciplining a learning theory as information is the most scarce at the point of labor force entry.

“shut down” distinction. Additionally, [Sorensen \(2016\)](#) shows workers separated according to a tenure-based rule system have lower earnings losses than those separated in less-restrictive manners. This is consistent with the employer learning theory because the presence of a “rule” works to eliminate selection based on individual fixed factors and so we would expect these workers to fare better.

4 Quantitative Analysis of Employer Learning

4.1 Calibration

The model predicts that observable wage dynamics following a separation to unemployment are informative about the unobservable process of employer learning. Accordingly, I calibrate parameter values to replicate the frequency of separations to unemployment, their serial correlation, and initial wage losses following unemployment. Non-targeted statistics include differences between selectively (endogenously) separated workers and those separated non-selectively (exogenously); as well as life-cycle wage growth and dispersion. The former set of statistics evaluates whether employer learning is quantitatively important for understanding wage dynamics after unemployment. The latter set asks the same question for the over-all wage process. I calibrate the model several times for a range of employer learning parameters—the speeds of incumbent learning and outside potential employer learning. In doing so, I provide a range of quantitative implications for employer learning speeds most supported by the Microeconomic employer learning literature.

The model period is one year.⁵⁰ I normalize the output of the low type worker $y_\ell = 1$. I choose a worker death rate of 0.025 to give an average working life of 40 years. I consider three speeds of incumbent learning (μ): 0.4, 0.5, 0.7. The middle value, 0.5 provides a best fit to the wage scar following unemployment and serves as the benchmark. The other two bounds provide a reasonable range of employer learning speeds found in the micro literature.⁵¹

⁵⁰Since there are no search frictions and potential employers immediately Bertrand compete for unemployed workers, the duration of unemployment is zero. Therefore, the assumption of a one-year duration is innocuous.

⁵¹For example, [Lange \(2007\)](#) estimates a public information model (no monopsony- wages equal expected

I consider two speeds of information diffusion to outside employers (ν):= 0.33 and 0.67. It is more difficult to align this parameter with microeconomic evidence because there is debate over whether employer learning is asymmetric. For this reason, I choose the lower value to be in line with [Kahn \(2013\)](#), who finds outside employers reduce their initial expectation errors by roughly a third of the incumbents’ reduction. The second higher value is in acknowledgement of studies that find outside employers learn at roughly the same rate, (e.g., [Schonberg \(2007\)](#)).⁵²

I maintain the PSID categories of “Shut down” and “Individually” separated workers defined in Section 3. I consider the former category to map accurately into a non-selective displacement in the model. However, it is reasonable to expect the latter category may include individuals separated individually, but not as a consequence of their own low-productivity.⁵³ Instead, I use the weaker assumption that the “individually separated” group contains more low-ability workers than the “shut-down” group. For this, I introduce a new parameter in the estimation: \hat{d} , the probability a worker is separated “exogenously”, but reports an “individual” separation in the PSID.

I rely on repeat displacement patterns predicted by the model to identify the share of individually separated workers in the PSID who were selectively terminated based on permanent low-ability. Assuming low-ability workers are more likely to report being “individually separated”, the proportion of individually separated workers separated again within the following year (FF) contains information about how many low-ability workers there are in the economy. For example, if the number experiencing individual separations two years in a row were close to the learning probability, μ , we would expect all separations were a consequence of employer learning. If this probability approached the total separation rate EU , we would expect all separations were exogenous, ie: independently and identically distributed across individuals.

productivity) and finds learning errors decline by 50% in 3 years. [Mansour \(2012\)](#) implements the same empirical strategy and finds learning can be as slow as 10 years for some occupations.

⁵²In the limiting case of symmetric learning $\nu = 1$, there would be no wage scars of unemployment in this model because the low-types would never be re-hired.

⁵³For example, their position within the firm could be eliminated because of the nature of the job, not the worker in it.

The final requirement is that the total separation rate must be consistent with the total empirical separation rate (EU). For a given μ , the proportion of high types π and exogenous separation hazard \hat{d} are calibrated to uniquely solve the following system of equations.⁵⁴

$$\begin{aligned} EU &= SD + \hat{d} + \mu(1 - \pi) \\ FF &= \frac{\hat{d}^2 + (\hat{d} + \mu)^2(1 - \pi)}{\hat{d} + \mu(1 - \pi)} \end{aligned}$$

Target values for the flow equation system are listed in Table 2. The total separation rate $EU = 0.051$ and the proportion of individually separated who are individually separated again within a year is 0.177.⁵⁵

This leaves two parameters to be determined: y_h : the productivity of the high type; and ρ , the discount rate. I choose ρ to equal one minus the probability of exogenous match destruction: the sum of exogenous separation hazard and the worker death rate. I choose y_h to match the estimated initial wage loss of separated workers upon re-employment. Specifically, I minimize the distance between the coefficient on a separation dummy for $t - 1$ in two Mincer regressions, one run on simulated data and the other on a PSID sample used in the prior section. *Crucially, both the model regression and the PSID regression include individual fixed effects, a quadratic in experience, and a full set of displacement dummies for 14 years following the event.*⁵⁶ Therefore, I am using truly comparable quantities as emphasized in [Barnette and Michaud \(2015\)](#).

4.2 Results

The complete list of parameters values and targets are shown in Table 3. The model was primarily calibrated to target separations to unemployment and the initial wage loss upon

⁵⁴Where the solution is unique by selecting only the positive root.

⁵⁵The proportion of workers individually separated a year after an exogenous shut down are not targeted, yet the model predictions are close to the data. Results for each calibration are shown in 4.

⁵⁶The PSID regression contains additional controls (urban area, industry, education, etc.) as detailed in the extended appendix. Obviously, these are not in the model regression.

reemployment. In this section I discuss the quantitative implications of the theory for *non-targeted* statistics.

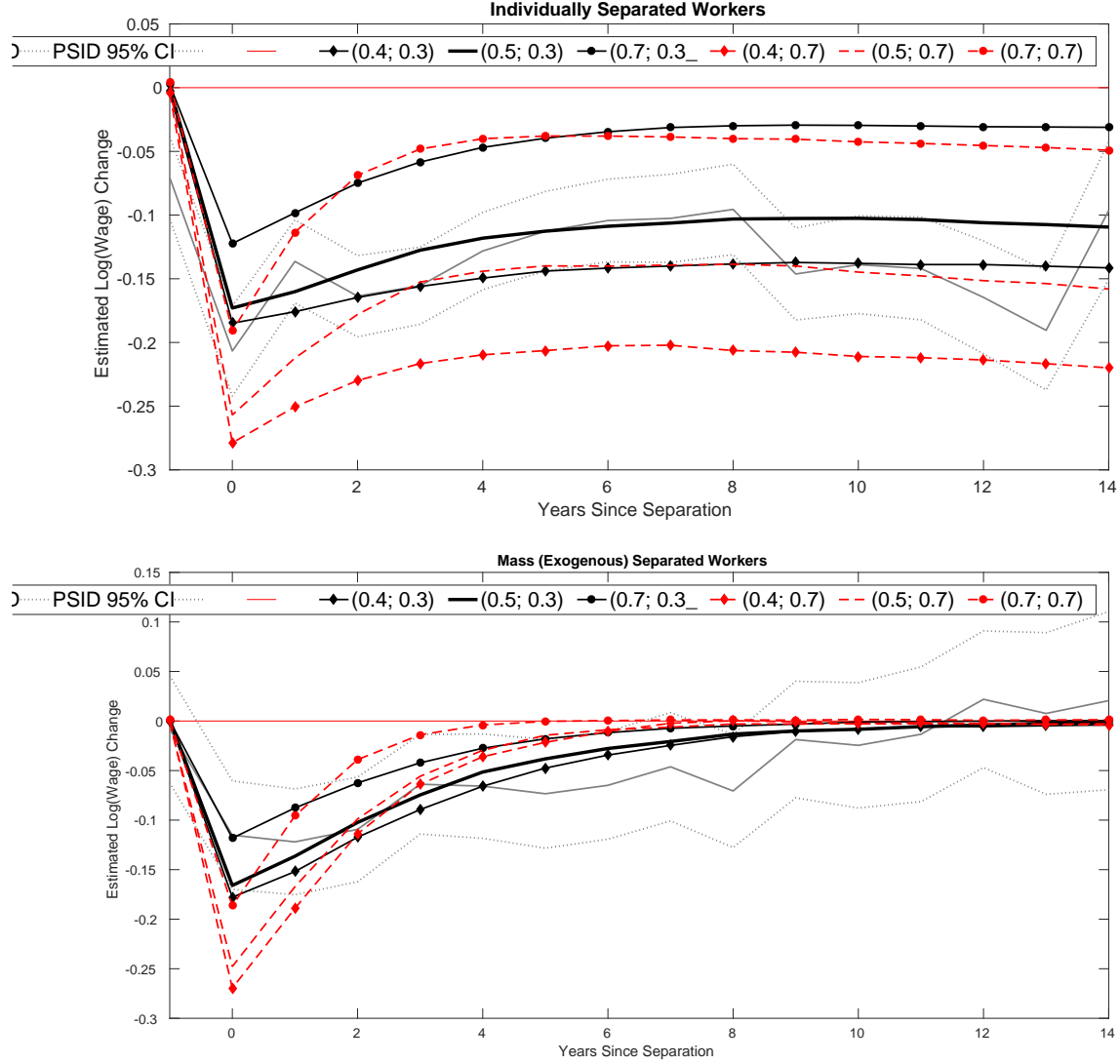


Figure 5: Regression coefficients of model and PSID data. Model versions listed as pairs (μ, ν) ; the speeds of current employer learning and of outside potential employer learning.

The first non-targeted aspect of model I evaluate is the ability to generate large, persistent wage losses of workers experiencing unemployment as well as differences between workers

separated for each reason. Graphical depictions are shown in Figure 5. The model replicates the divergence in paths across shut down and individually separated workers: the shut down workers recover and the individually separated workers do not. This is one success of the model because the calibration does not target the dynamics of shut-down workers and fired workers separately. These results were driven entirely by the theory that some individually fired workers are selected because of low, permanent ability. Quantitatively, the baseline calibration accounts for 77% of the present discounted value of the wage losses of individually separated workers and 126% of the wage losses of Shut-Down workers.⁵⁷ While the former (latter) is slightly under-stated (over-stated), the fact that the paths of each type of worker diverge based on reason for unemployment is unique to the theory developed in this paper.

The third set of statistics deals with life-cycle wage growth and cross-sectional dispersion. The calibrations range from generating 16-42% of life-cycle wage growth and 8.9-20% of the standard deviation of residual wages after 25 years of experience.⁵⁸ The baseline calibration generates values of 24% and 13%, respectively. These values are modest, but not unreasonable given the model does not have any changes in actual worker productivity over the life-cycle.

4.3 Comparative Statics on Learning Speeds

A distinction of this work is that it is flexible in the speed of employer learning and the speed of diffusion of information to outside potential employers.⁵⁹ In this section I provide comparative statics on learning speeds and other key parameters of the model to better understand what mechanisms deliver the quantitative implications of employer learning and their robustness. I begin with analytic propositions on the channels through which parameters affect outcomes

⁵⁷I use a discount of $\beta = .95$ just as Davis and von Wachter (2011) and calculate the ratio of the sum of the coefficients on the 14 displacement dummies in the model regression divided by the same sum in the data regression. For comparability, the empirical estimates of wage losses from unemployment in their sample is 1.4-2.8 years of earnings over 20 years. The estimates from my sample are 1.53 years of earnings over 14 years, but do not include periods of zero earnings.

⁵⁸Calculated as ratio of the equally weighted sum of a rolling 5 year window of median log wages (or standard deviations of residuals) over age 30-55 in the model divided by the same statistic in the data.

⁵⁹Two period models implicitly assume employers learn after one period, $\mu = 1$ in my model. Models of private learning assume poachers receive no information $\nu = 0$ and public learning models assume poachers receive all information $\nu = 1$.

and then present quantitative comparative statics on the baseline calibrated model.

Wages offered to new entrants and the unemployed each satisfy Bertrand competition and are affected by the learning speeds through expected future profits. As shown in Equation 2.5 and Equation 2.4, the expected profits depend on three factors. The first is the expected output of the current worker $y(p)$. For new entrants, this depends on the proportion of high types in the economy. For workers hired from unemployment, this depends on the proportion of unemployed that are high type: $\frac{d\pi}{d+\mu(1-\pi)}$. By inspection, this proportion is increasing in exogenous flows to unemployment d and the proportion of high-types in the economy π . It is decreasing in the current employer learning speed μ , which increases the selective separation hazard for low-types. The second factor is how long it takes to terminate a match with a low-type. The faster the speed of current employer learning μ , the faster low-types will be terminated and the higher expected profits will be. The final factor is the duration an employer gets to keep a high-ability worker after learning their type. In this situation, the employer earns positive ex-post profits because they need not raise the worker's wage until potential employers learn her type. Thus, expected profits are decreasing in both the speed of outside potential employer learning (ν is small) and the exogenous separation rate δ .

Proposition 4.1 states the effect of parameters on the share of high-types p^u in the stationary distribution of the unemployed.⁶⁰ Proposition 4.2 formalizes the effect of the latter two channels on wages, holding the first channel (composition of unemployed) fixed. In some cases these channels move wages in opposite directions leading to ambiguous results. A quantitative analysis on a calibrated model follows

Proposition 4.1. *The proportion of unemployed that are high-types is (i) increasing in π - the proportion of high-types in the economy; (ii) decreasing in μ - the Poisson arrival of incumbent learning; (iii) unaffected by ν - the Poisson arrival of informed poachers; (iv) increasing in d - the exogenous separation rate.*

$$\frac{dp^u}{d\pi} > 0 \quad \frac{dp^u}{d\mu} < 0 \quad \frac{dp^u}{d\nu} = 0 \quad \frac{dp^u}{dd} > 0$$

⁶⁰The wage offered to an unemployed worker is trivially increasing in p^u by Bertrand Competition.

Proof. by inspection on $p^u = \frac{d\pi}{d+\mu(1-\pi)}$. □

Proposition 4.2. *Holding fixed rational priors p , wage offers to new entrants and the unemployed are each: (i) increasing in μ - the Poisson arrival of incumbent learning; (ii) decreasing in ν - the Poisson arrival of informed poachers; (iii) decreasing in d - the exogenous separation rate.*

$$\frac{dw^*(p)}{d\mu} > 0 \quad \frac{dw^*(p)}{d\nu} < 0 \quad \frac{dw^*(p)}{dd} < 0$$

Proof. by inspection on Equations 2.5 and 2.4. □

An interesting corollary is that the wage paid to each new entrants and those hired from unemployment is strictly larger than the expected per-period output of a worker from each pool, so long as $\nu < 1$. This is because Bertrand Competition is forward looking. Low-types will yield a negative profit each period, but they will be terminated sooner than high-types who yield a positive profit each period.⁶¹

Corollary 4.3. *Wage offers to new entrants and the unemployed are higher than their expected per-period output.*

$$w^*(p^{ne}) > y(p^{ne}) \quad \text{and} \quad w^*(p^u) > y(p^u)$$

Proof. Omitted. □

Comparing Propositions 4.1 and 4.2, one can see that some parameters such as ν have unambiguous effects on wages paid to new entrants and the unemployed while other parameters such as μ and d have ambiguous effects: the impact on the composition of the pool of unemployed works in the opposite direction as the impact of these parameters on the expected match duration of high-types relative to low. I now provide comparative statics on the baseline calibration to measure the quantitative outcome of these competing mechanisms. The baseline calibration is for $\mu = 0.5$ and $\nu = 0.33$. I fix parameters calibrated under these values $(\hat{d}, \rho, \pi, y^h)$ and consider 6 unilateral departures from the baseline for each of: $\mu = \{0.2, 0.4, 0.8\}$ and $\nu = \{0.2, 0.6, 0.8\}$. For each departure I calculate model predictions

⁶¹This is a departure from empirical models that assume wages are equal to the expected marginal product of a worker.

for wage scars, life-cycle wage growth, and cross-sectional dispersion. Results are provided graphically in Figure 6 and Table 6.

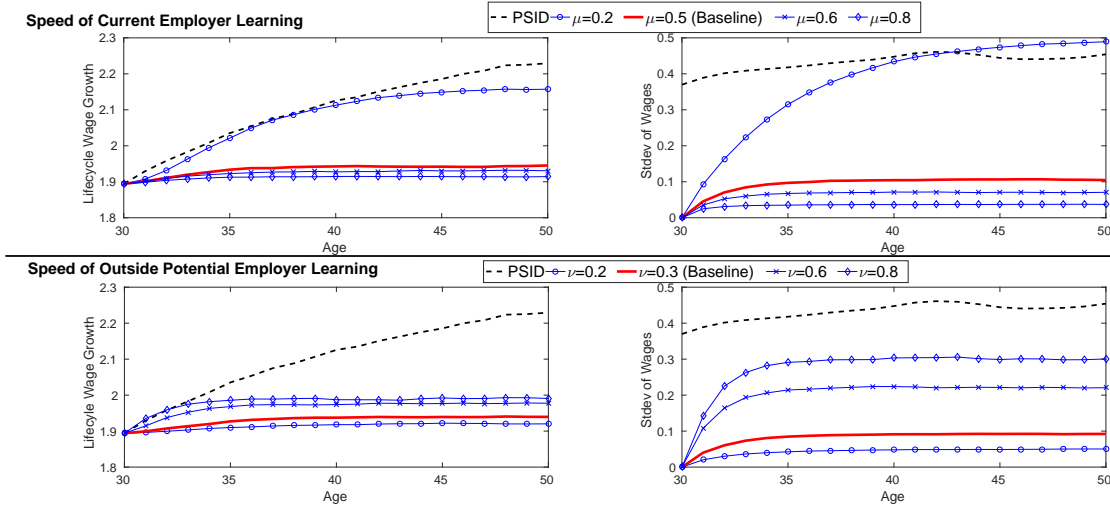


Figure 6: Comparative Statics: Speed of Learning on Wage Growth and Dispersion

Speed of Incumbent Employer Learning (μ). The upper left pane of Figure 6 shows that slower current employer learning provides greater life-cycle wage growth. This is because a slower speed of employer learning lowers the initial wage offered to new entrants by increasing the time it takes for an employer to discover and terminate a low-ability type (Proposition 4.1).⁶² For similar reasons, a slower speed of current employer learning increases cross-sectional wage dispersion over the life-cycle. Workers experiencing unemployment have even larger wage losses when learning is slow because the negative effect of longer time to discover a low-ability type on expected profits (Proposition 4.1) outweighs the positive effect of fewer low-ability workers in the unemployment pool (Proposition 4.2) as shown in Table 6.

Speed of Diffusion of Information to Potential Poaching Employers (ν). Faster outside learning unambiguously decreases the wages offered to both new entrants and unemployed (Proposition 4.2) and increases the speed of wage growth for high-ability types. It

⁶²I normalize all starting wages to equal the PSID log wage value in this figure.

takes less time for poachers to learn a worker’s type and offer her the full match surplus. These effects work together to provide larger life-cycle wage growth and cross-sectional wage dispersion as outside potential employer learning becomes faster. Wage scars are increased for the same reason: faster outside learning increases the wages of the reference group to which separated workers are compared.

Other parameters. A greater proportion of high ability types in the economy (π) raises each the new entrant wage and the wage of a worker hired from unemployment. Impact on wage scars and overall wage growth is non-monotone. At both $\pi = 0$ and $\pi = 1$, wage growth is zero and wage scars in unemployment are zero. A greater exogenous separation (d) rate lowers the wage paid to new entrants and has an ambiguous affect on the wage paid to a worker hired from unemployment. The reasoning is similar to the discussion of μ . A higher exogenous separation rate increases wages offered to unemployed through Proposition 4.1, but decreases it through 4.2. Intuition is found in the extremes. At an exogenous separation rate approaching one unemployed workers are treated identically to new entrants and there are no wage scars. At an exogenous separation rate approaching zero, the wage scars are the largest because the composition of the unemployed becomes entirely selectively separated low-types.

5 Conclusion

This paper evaluated the role of employer learning in the process coupling wage dynamics with employment transitions. I have shown this theory predicts diverging wage and employment patterns of separated and never separated workers; as well as key differences within the latter group by the reason for their separation. I verify these predictions in the PSID. I find large and persistent wage losses coupled with repeat unemployment for selectively separated workers alongside the recovery of workers separated at no fault of their own. I then evaluate whether employer learning is an empirically significant part of wage dynamics and employment flows. In the baseline calibrated model, I find employer learning accounts for 78% of average wage losses following unemployment and generates double the wage losses for selectively separated

workers. The theory of employer learning also accounts for 24% of life-cycle wage growth and 13% of cross-sectional dispersion.

The results in this paper suggest employer learning is an important factor in understanding worker turnover and wage dynamics, particularly for workers moving through unemployment. This is interesting because the employer learning mechanism generates wage changes without any change in productivity. Most striking is the stigma effect of unemployment that delivers wage losses for highly-productive separated workers that persist until their reputation is recovered. This motivates a desire to revisit research questions dealing with the productivity and welfare implications of unemployment, as well as policy prescriptions based on these analyses. The model I have constructed is useful for this end: to study the quantitative implications of employer learning related to both positive and normative questions. Firm's profits, wages and worker turnover are all endogenously linked in a tractable dynamic model. This opens the opportunity for job creation and equilibrium unemployment to be endogenized in a search framework.

A caveat is that the assumptions implemented in this paper may amplify the importance of employer learning. For example, potential employers do not observe workers' job tenure or unemployment duration. The Poisson learning structure of this model can accommodate these elements while maintaining tractability. However, incorporating match specific or employer heterogeneity remains a non-trivial extension left as a challenge for future research.⁶³

References

ACEMOGLU, D. AND J. PISCHKI (1996): "Why Do Firms Train? Theory and Evidence," Working papers 96-7, Massachusetts Institute of Technology (MIT), Department of Economics.

⁶³There are many features common to labor market models that are omitted from the analysis. Including additional features complicates the Bayesian posteriors of employers (beliefs about the worker conditional on the fact she accepts the offer) by expanding the set of variables they condition on. Including match-specific components whether pecuniary (Jovanovic (1979)) or non (Acemoglu and Pischki (1996)) compromises the pooling equilibrium. They make it potentially costly for the low-ability type to mimic the high type, opening the opportunity for separating contracts.

- AKERLOF, G. A. (1970): “The Market for Lemons: Quality Uncertainty and the Market Mechanism,” *The Quarterly Journal of Economics*, 84, 488–500.
- ALTONJI, J. G. AND C. R. PIERRET (2001): “Employer Learning And Statistical Discrimination,” *The Quarterly Journal of Economics*, 116, 313–350.
- BARNETTE, J. AND A. M. MICHAUD (2015): “Wage Scars from Job Loss,” *Mimeo*.
- BECKER, G. (1975): *Human Capital*, Columbia University Press, 2 ed.
- BEN-PORATH, Y. (1967): “The production of human capital and the life cycle of earnings,” *Journal of political economy*, 75, 352–365.
- BURDETT, K., C. CARRILLO-TUDELA, M. G. COLES, ET AL. (2015): “The Cost of Job Loss,” Tech. rep., Institute for Social and Economic Research.
- BURDETT, K. AND D. T. MORTENSEN (1998): “Wage differentials, employer size, and unemployment,” *International Economic Review*, 257–273.
- COUCH, K. A. AND D. W. PLACZEK (2010): “Earnings Losses of Displaced Workers Revisited,” *American Economic Review*, 100, 572–89.
- DAVIS, S. J. AND T. M. VON WACHTER (2011): “Recessions and the cost of job loss,” Tech. rep., National Bureau of Economic Research.
- FLAAEN, A., M. D. SHAPIRO, AND I. SORKIN (2013): “Reconsidering the consequences of worker displacements: Survey versus administrative measurements,” Tech. rep., Working paper, University of Michigan, September 5.
- GIBBONS, R. AND L. F. KATZ (1991): “Layoffs and Lemons,” *Journal of Labor Economics*, 9, 351–80.
- GUVENEN, F., B. KURUSCU, S. TANAKA, AND D. WICZER (2015): “Multidimensional skill mismatch,” Tech. rep., National Bureau of Economic Research.
- HARRIS, M. AND B. HOLSTROM (1982): “A Theory of Wage Dynamics,” *Review of Economic Studies*, 49, 315–33.
- HECKMAN, J. J. (1976): “A Life-Cycle Model of Earnings, Learning, and Consumption,” *Journal of Political Economy*, 84, S9–S44.
- JACOBSON, L. S., R. J. LALONDE, AND D. G. SULLIVAN (1993): “Earnings losses of displaced workers,” *The American Economic Review*, 685–709.
- JAROSCH, G. (2014): “Searching for job security and the consequences of job loss,” *V University of Chicago*,

Job Market Paper.

- JOVANOVIĆ, B. (1979): “Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 87, 972–90.
- JUNG, P. AND M. KUHN (2016): “Earnings losses and labor mobility over the lifecycle,” Tech. rep., CEPR Discussion Papers.
- KAHN, L. B. (2013): “Asymmetric information between employers,” *American Economic Journal: Applied Economics*, 5, 165–205.
- KREBS, T. (2007): “Job displacement risk and the cost of business cycles,” *The American economic review*, 664–686.
- KROLIKOWSKI, P. (2017): “Job ladders and earnings of displaced workers,” *American Economic Journal: Macroeconomics*, 9, 1–31.
- LANGE, F. (2007): “The Speed of Employer Learning,” *Journal of Labor Economics*, 25, 1–35.
- LAZEAR, E. P. AND P. OYER (2007): “Personnel economics,” Tech. rep., National Bureau of economic research.
- LAZEAR, E. P. AND S. ROSEN (1981): “Rank-order tournaments as optimum labor contracts,” *Journal of political Economy*, 89, 841–864.
- LISE, J. AND F. POSTEL-VINAY (2015): “Multidimensional skills, sorting, and human capital accumulation,” Tech. rep., Working Paper, 2015. University College London.
- LJUNGQVIST, L. AND T. J. SARGENT (1998): “The European Unemployment Dilemma,” *Journal of Political Economy*, 106, 514–550.
- LOCKWOOD, B. (1991): “Information Externalities in the Labour Market and the Duration of Unemployment,” *The Review of Economic Studies*, 58, 733–753.
- MANSOUR, H. (2012): “Does employer learning vary by occupation?” *Journal of Labor Economics*, 30, 415–444.
- MINCER, J. A. (1974): “Schooling and earnings,” in *Schooling, experience, and earnings*, NBER, 41–63.
- NICKELL, S. J. (1976): “Wage structures and quit rates,” *International Economic Review*, 191–203.
- PALLAIS, A. (2014): “Inefficient hiring in entry-level labor markets,” *American Economic Review*, 104, 3565–3599.
- PINKSTON, J. C. (2009): “A Model of Asymmetric Employer Learning with Testable Implications,” *Review*

of *Economic Studies*, 76, 367–394.

POSTEL-VINAY, F. AND J.-M. ROBIN (2002): “Equilibrium wage dispersion with worker and employer heterogeneity,” *Econometrica*, 70, 2295–2350.

SCHONBERG, U. (2007): “Testing for Asymmetric Employer Learning,” *Journal of Labor Economics*, 25, 651–691.

SORENSEN, J. (2016): “Layoff rules, the cost of job loss, and asymmetric employer learning,” *UC Berkely-Job Market Paper*.

STEVENS, A. H. (1997): “Persistent Effects of Job Displacement: The Importance of Multiple Job Losses,” *Journal of Labor Economics*, 15, 165–88.

WALDMAN, M. (1984): “Job assignments, signalling, and efficiency,” *The RAND Journal of Economics*, 15, 255–267.

6 Appendix

Table 1: Logistic Regression on Separation Hazards: Marginal Effects and Hypothesis Tests

Marginal Effects		
Indep. Variables	Marginal Effect; (s.e.)	
	(1) Individual t	(2) Shut Down t
Individual Separation $t - 1$ (β_{t-1}^F)	0.0362 **; (0.0025)	0.0086 **; (0.002)
Shut Down (Exog) Separation $t - 1$ (β_{t-1}^{SD})	-0.0001; (0.985)	0.0146 ; (0.002)
Hypothesis Testing		
H0	chi2(1); (p-value)	
	(1) Individual t	(2) Shut Down t
$\beta_{t-1}^F = 0$	274.56**; (0.00)	30.25**; (0.00)
$\beta_{t-1}^{SD} = 0$	0.00; (0.985)	63.16**; (0.00)
$\beta_{t-1}^F = \beta_{t-1}^{SD}$	22.31**; (0.00)	6.62*; (0.01)

Significance levels : † : 10%; * : 5%; ** : 1%

Table 2: Flow Equation System: PSID Targets and Model Fit

Statistic	Target	Model Value (for all calibrations)
Annual Proportion Fired or shut down	0.051	0.051
Prop. fired yr after a firing	0.177	0.177
Prop. fired yr after Shut Down (non-targeted)	0.044	see Table 4

In all versions parameter values are exactly identified by a system of flow equations.)

Table 3: Baseline and Alternative Calibrations: Parameter Values

Incumbent Learning Speed μ	0.4	0.5	0.7	0.4	0.5	0.7
Poacher Learning Speed ν	0.33			0.67		
Exogen. Separation Hazard $d + \hat{d}$	0.009	0.010	0.015	0.009	0.010	0.015
High Types π	0.939	0.942	0.949	0.939	0.942	0.949
High Type Output y^H	2.00	2.03	2.01	2.02	2.03	2.04
Discount Rate ρ	0.992	0.990	0.985	0.991	0.990	0.985

Baseline calibration in bold (learning speeds: $\mu = 0.5, \nu = 0.33$).

Table 4: Baseline and Alternative Calibrations: Model Implications (Non-Targeted)

Statistic	PSID	Model					
Incumbent Learning Speed μ		0.4	0.5	0.7	0.4	0.5	0.7
Poacher Learning Speed ν		0.33			0.67		
Coef Year after Select Separation	-0.17	-0.19	-0.17	-0.12	-0.28	-0.26	-0.19
pdv wage loss after Select Sep.	-1.57	-1.19	-0.95	-0.45	-1.68	-1.30	-0.55
Coef Year after Exog Separation	-0.12	-0.14	-0.10	-0.03	-0.21	-0.15	-0.04
pdv wage loss after Exog Sep.	-0.64	-0.98	-0.78	-0.37	-1.28	-1.01	-0.43
% Select Sep. yr after Exog Sep.	0.044	0.027	0.032	0.044	0.027	0.032	0.044
% of Separations that are Selective	n/a	84.0	80.4	71.6	84.0	80.4	71.6
20 yr wage growth (average)	1.74	1.19	1.18	1.12	1.31	1.28	1.20
20 yr wage dispersion (stdev)	0.45	0.07	0.06	0.04	0.09	0.09	0.05

Baseline calibration in bold (learning speeds: $\mu = 0.5, \nu = 0.33$). Model re-calibrated for additional values of μ and ν for robustness.

Table 5: Baseline and Alternative Calibrations: Accounting for PSID Facts

Statistic	PSID	Model					
Incumbent Learning Speed μ		0.4	0.5	0.7	0.4	0.5	0.7
Poacher Learning Speed ν		0.33			0.67		
Short term loss: Select Separation	-0.17	112%	100%	71%	165%	153%	112%
pdv wage loss after Select Sep.	-1.57	76%	61%	29%	107%	83%	35%
Short term loss: Exog Separation	-0.12	117%	83%	25%	175%	125%	33%
pdv wage loss after Exog Sep.	-0.64	153%	122%	58%	200%	158%	67%
20 yr wage growth (average)	1.74	26%	24%	16%	42%	38%	27%
20 yr wage dispersion (stdev)	0.45	16%	13%	8.9%	20%	17%	11%

Percentage of PSID statistic accounted for by employer learning. Numbers greater than 100% mean the model overstates the statistic.

Short term loss measured as the average coefficient in years 0-2.

Table 6: Comparative Statics on Baseline Calibration

Statistic	Baseline	Comparative Statics					
Incumbent Learning Speed μ	0.5	0.2	0.6	0.8	0.5		
Poacher Learning Speed ν	0.33	0.33			0.2	0.6	0.8
pdv wage loss after Select Sep.	-0.95	-2.38	-1.73	-0.86	-0.68	-2.31	-2.81
pdv wage loss after Exog Sep.	-0.78	-2.25	-1.54	-0.74	-0.43	-1.51	-1.74
20 yr wage growth (average)	1.18	1.68	1.16	1.12	1.12	1.28	1.30
20 yr wage dispersion (stdev)	0.06	0.49	0.05	0.03	0.03	0.21	0.30

Baseline calibration is for $\mu = 0.5, \nu = 0.33$. This table shows impact of single parameter changes.