Replacement Hiring and the Productivity-Wage Gap *

Sushant Acharya
FRB New York

Shu Lin Wee
Carnegie Mellon University

October 4, 2019

Abstract

A large and growing share of hires in the US are replacement hires. This increase coincides with a growing productivity-wage gap. We connect these trends by building a model where firms post long-lived vacancies and both workers and firms engage in on-the-job search for more productive matches. Unlike worker on-the-job search which causes wages to grow at a faster rate with improvements in match quality, increased job insecurity and lower matching efficiency from increased firm on-the-job search reduce workers' outside options, causing wages to grow at a slower rate and prompting the emergence of a productivity-wage gap. Quantitatively, the model can account for the observed widening of the productivity-wage gap.

Keywords: Replacement hiring, Productivity-wage gap, Unemployment, Labor share, Efficiency
JEL Codes: E32, J63, J64

*The authors would like to thank Isabel Cairo, Keshav Dogra, Ross Doppelt, John Haltiwanger, Ryan Michaels and Laura Pilossoph for useful comments and suggestions. We also thank the seminar/conference participants at the Census Bureau, Carnegie Mellon University, EIEF, University of Edinburgh, FRB Cleveland, FRB Kansas City-CADRE Workshop, FRB St. Louis, MidWest Macro Pittsburgh, 9th European Sam Network Annual conference. This paper was formerly titled "Replacement Hiring". The views expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the Federal Reserve Bank of New York or the Federal Reserve System.
Email: Acharya: sushant.acharya@ny.frb.org , Wee: shuwee@andrew.cmu.edu

1 Introduction

The last two decades have seen an increasing divergence between labor productivity and wages. Figure 1 shows that prior to 2000, real compensation per hour grew at roughly the same rate as real output per hour, i.e. labor productivity.[1] Post 2000, however, there has emerged a divergence between labor productivity and wage compensation. This gap between labor productivity and wage compensation - which we term the productivity-wage gap - continues to grow even as the unemployment rate continues to reach new lows in the aftermath of the Great Recession and the quits rate has surpassed its pre-recession level, suggesting that the increase in the gap is not due to slack labor markets. Given this backdrop, we ask instead how the incidence of firm on-the-job search and its impact on outside options can affect the productivity-wage gap.

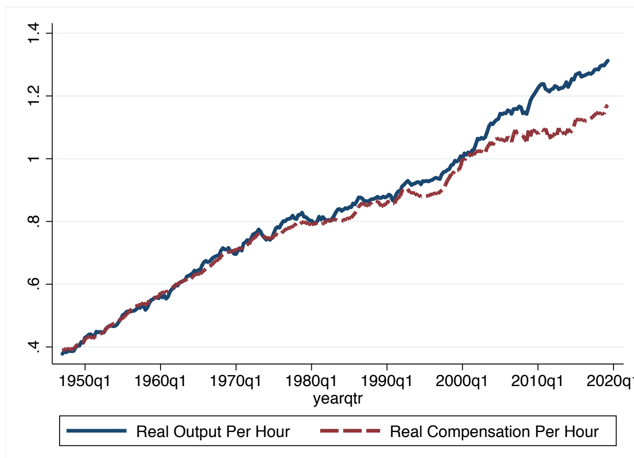


Figure 1: Labor productivity and real hourly compensation

Notes: (i) Data on Current Output Per Hour and Current Compensation per Hour comes from the U.S. Bureau of Labor Statistics Productivity and Costs database. We deflate both measures using the consumer price index (CPI).

The way firm on-the-job search manifests itself is through replacement hiring - firms who seek higher quality applicants replace current workers with better workers. Importantly, over the same time period, there has been an upward trend in the fraction of total hires that are replacement hires.[2] Replacement hires are defined by the Census as hires that continue into the next period in excess of net employment change.[3] Using data from the Quarterly Workforce Indicators (QWI), Figure 2a shows that the fraction of total hires that are replacement hires has increased from about 33% in the early 1990s to a high of about 41% in 2017.[4] While the time aggregation at a quarterly frequency implies that replacement hires are also recorded whenever a firm re-fills a vacated position, Figure 2b shows that employment-

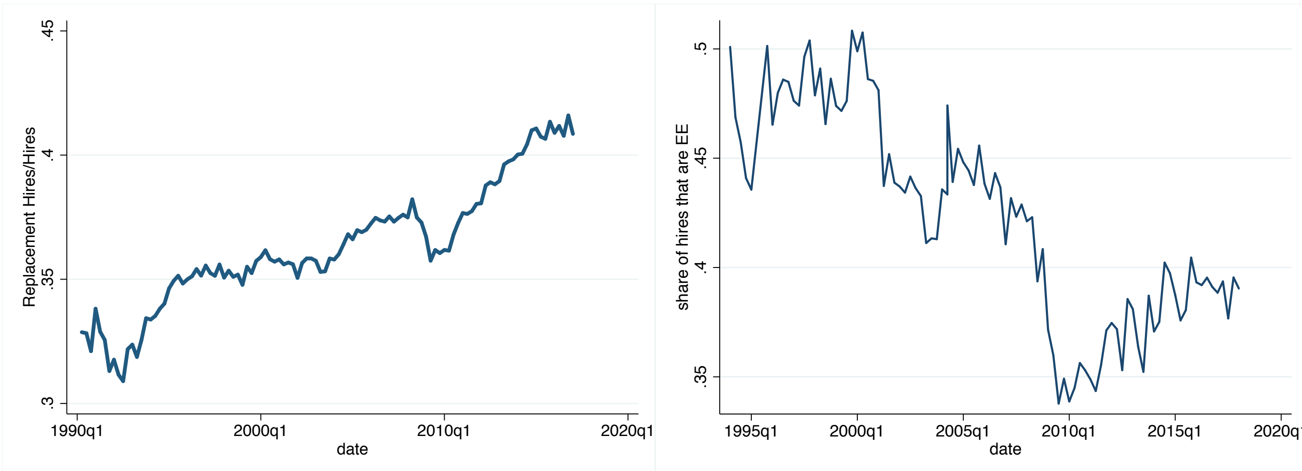
[1]To calculate real compensation per hour and real output per hour, we take current output per hour and current compensation per hour and deflate both measures using the CPI. Note that by using a common price index, the divergence in the productivity and wages stems not a difference in price deflators.

[2]It is important to note that the increased share of replacement hires is not inconsistent with declining labor mobility and declining trends in job creation. In fact, as a fraction of average total employed, the replacement hiring *rate* has been declining over time. Figure 7a in Appendix A shows that the hiring rate has fallen faster than the replacement hiring rate. The sharper decline in total hires relative to replacement hires implies that replacement hiring is increasingly becoming a more important share of total hiring.

[3]See Section 2 for more details. All definitions are taken from `https://lehd.ces.census.gov/doc/QWI_101.pdf`.

[4]Information collected on replacement hiring as recorded in the QWI only begins from the 1990s.

to-employment transitions (EE hires) as a fraction of total hires has trended downwards.[5] The rise in the replacement hiring share amid the decline in the ratio of EE hires to total hires suggests that firm on-the-job search may be an overlooked channel which is growing in importance.



(a) Replacement hires as a share of total hires

(b) EE hires/total hires

Figure 2: Trends in replacement hiring vs. worker job-to-job transitions

Notes: (i) The replacement hiring share is the ratio of replacement hires to hires. We formalize this definition in Section 2. (ii) The EE hires share is the number of employed individuals who moved to a new employer divided by hires. This measure is calculated using data from the Current Population Survey (CPS). To capture the numerator of this measure, we employ the same techniques as in Fallick and Fleischman (2004).

We view firm on-the-job search as a natural starting point for understanding how a wider productivity-wage gap could emerge. To this end, we build a model that features both worker and firm on-the-job search. In our model, search is random and firms pay a fixed cost to create a new vacancy. A vacancy in our model is synonymous with a job position being created. A vacancy or job position is long-lived and is not destroyed instantaneously. Rather, a vacancy continues to exist even if the firm fails to fill it immediately. Further, firms whose job positions have been filled can continue to meet applicants i.e, firms can conduct on-the-job search, so long as the job position has not been destroyed. When firms are allowed to search on-the-job, they seek applicants who are better matches than their current workers, and who can bring higher output to the firm. In the same vein, workers in our model also conduct on-the-job search so as to meet vacancies who are better matches than their current firms. In their efforts to match with higher productivity applicants and firms, both firm and worker on-the-job search cause labor productivity to increase. Thus, an increased incidence of either worker or firm on-the-job search is associated with higher labor productivity.

Worker and firm on-the-job search however, have different implications for wages. While increased worker on-the-job search raises the worker's value of employment as well as their value of unemployment, i.e. their outside option, firm on-the-job search has the opposite implication. Here, the ability of firms to conduct on-the-job search changes the outside options of both firms and workers. By extending the standard search model to allow for long-lived vacancies and firm on-the-job search, we uncover three channels through which firm on-the-job search can depress wages while still raising productivity. First, the fact that vacancies are long-lived imply a positive option value to holding a vacancy. This positive

[5] We further show in Section 2 that not all of replacement hiring in the data can be accounted for by quits.

option value or *market power* raises a firm's outside option when bargaining with the worker, allowing the firm to keep wages low and extract a larger share of the surplus. Secondly, firms' ability to conduct on-the-job search raises workers' effective separation rate and the amount of *job insecurity* they face. Increased job insecurity in turn reduces the average employment spell and diminishes workers' outside options, further allowing firms to pay lower wages. Finally, when firms can search on-the-job, the composition of vacancies now comprises of both unfilled vacancies as well as currently matched firms with unexpired vacancies - which we label as *recruiting matched firms.* For an unemployed applicant to be hired at this latter type of vacancy, her productivity must be higher than that of the firm's incumbent worker. Thus, relative to meeting an unfilled vacancy, unemployed job-seekers must pass a higher bar before they are hired by the firm. This generates a larger wedge between hiring and meeting rates - lowering *measured matching efficiency*, which in turn diminishes workers' outside options and further reduces their wages.

To investigate the impact of firm vs. worker on-the-job on the productivity-wage gap, we conduct two separate exercises. First, we provide some comparative statics. Holding all else constant, we show that an increase in worker's (firm's) ability to conduct on-the-job search leads to a productivity distribution that first order stochastically dominates a distribution observed under the case when worker's (firm's) ability to conduct on-the-job search is low. As such, both firm and worker on-the-job search lead the productivity distribution to exhibit larger mass at higher productivity values, thus causing average labor productivity to increase. The rate at which wages grow with this improvement in productivity, however, depends on whether worker or firm on-the-job search is more prevalent. We show in our comparative static exercises, that the slope of the wage function with respect to productivity is steeper when worker on-the-job search is high. In contrast, the rate at which wages grow with productivity is lower when firm on-the-job search is high.

Second, we calibrate the model to match labor market flows in the US as well as the replacement hiring share and EE hiring share across two time periods: before 2000 and after 2000. We use the year 2000 as our cut-off as the data shows that the productivity-wage gap only started to diverge post 2000. In conducting this exercise, we examine if our model, when calibrated to match the rise in the replacement hiring share, can at the same time replicate the observed wider gap between productivity and wages. In our quantitative exercise, we measure replacement hiring as having occurred whenever firms search on-the-job and when they refill a vacated position. Across the time periods, our model predicts that the higher replacement hiring share widened the productivity-wage gap by 12%, close to the total increase of 8% in the productivity-wage gap over the time periods of interest.[6].

We further use our calibrated model across these two time periods to identify the role of worker vs. firm on-the-job search towards the widening productivity wage gap. In our counterfactual exercises, we vary the parameter governing the opportunity with which workers (firms) could conduct on-the-job search while keeping all other parameters constant at their post 2000. We find that the change in the parameter governing the opportunity for workers to do on-the-job search did little to change the productivity-wage gap. In contrast, keeping the parameter that governed the opportunity for firms to

[6]We measure the productivity-wage gap as the ratio of labor productivity to mean compensation. Since the our data for replacement hiring only begins in the 1990s, we only calculate the change in the productivity-wage gap between the periods 1990-2000 and post 2000.

do on-the-job search at its higher pre-2000 level would have further widened the productivity-wage gap. Our results underscore the role of firm on-the-job search in affecting the productivity-wage gap.

Related Literature Our paper contributes to the growing literature on the impact of replacement hiring. Two papers are closely related to ours. Both Mercan and Schoefer (2019) and Elsby et al. (2016) focus on the business cycle properties of vacancy posting and examine how replacement hiring and quits by workers necessitates a firm to re-fill a position. Importantly, the two aforementioned papers focus on worker on-the-job search while we examine the implications of on-the-job search by firms and workers on the productivity-wage gap. Separately, Menzio and Moen (2010) examine replacement hiring in the context that firms seek to insure workers against income fluctuations but cannot commit to not replacing current workers in a downturn with cheaper new hires. While their paper is concerned with characterizing the efficient wage contract, we examine how, even absent aggregate shocks, the decentralized economy can create too many low productivity matches relative to what a planner would choose . In related work, Kiyotaki and Lagos (2007) study a random search model which features both replacement hiring and worker on-the-job search. They examine the extent to which the decentralized economy can implement the planner's outcome when workers and firm both engage in Bertrand competition in terms of life-time utility offers. Importantly, their paper abstracts from wages while our paper is primarily concerned with how firm on-the-job search can widen the productivity-wage gap through replacement hiring.

Our paper is also related to the literature on long-lived vacancies. Fujita and Ramey (2007) and Haefke and Reiter (2017) consider models where job positions are long-lived and firms do not shut down immediately upon worker separation. Both of these papers demonstrate that the inclusion of long-lived vacancies in a labor search model can better replicate labor market flows in the data. Firms with unexpired job positions in these models only re-hire new workers when they are separated from their current worker. As such, these models do not address the issue of firm on-the-job search and its ramifications for the productivity-wage gap.

Although we study how firm on-the-job search can affect the productivity-wage gap, our paper also has implications for the labor share. Intuitively, the divergence in labor productivity and compensation implies that a smaller share of total output accrues to labor. Karabarbounis and Neiman (2013) document that the labor share has declined across countries and argue that capital deepening is the primary factor behind this decline. Elsby et al. (2013) conduct a comprehensive study and find a strong negative relationship between import exposure and the labor share at the industry level. We add to this debate by showing how firm on-the-job search can lead to a smaller labor share by raising productivity while keeping wages low. Recent work by Autor et al. (2017) and Azar et al. (2017) suggest that product market concentration is associated with labor market concentration. These papers empirically show that wages are lower when firms observe increased market power. By focusing on long-lived vacancies, our paper offers an alternative view of firm market power and shows how allowing for firm on-the-job search depresses the outside option of workers relative to firms.

Finally, our paper is also related to the recent literature on phantom vacancies. Cheron and Decreuse (2017) and Albrecht et al. (2017) argue that phantoms are vacancies that have already found a match and cannot generate any more new hires. The existence of phantoms lowers matching efficiecy as unemployed jobseekers cannot convert a meeting with a phantom into a hire. We offer an alternative view: matched firms with unexpired vacancies can still generate hires. An unemployed job applicant who contacts a recruiting matched firm, however, must surpass the productivity of the incumbent worker before she is hired. As such, these long-lived vacancies which allow firms to conduct on-the-job search also lowers measured matching efficiency. Recent work using online vacancy job board data by Davis and de la Parra (2017) suggests that a non-trivial portion of job postings are "long-duration" postings which are continuously on the look-out for new applicants, giving support to our supposition that vacancies are long-lived and can re-match with multiple workers.

The rest of this paper is organized as follows. Section 2 discusses the data on replacement hiring. Section 3 introduces the model. Section 4 outlines our comparative static exercises and highlights how worker vs. firm on-the-job search affects wages and productivity. Section 5 explores the predictions of our model while Section 6 contains a brief discussion about some assumptions of our model Finally, Section 7 concludes.

2 Data

Building on the underlying Longitudinal Employer Household Dynamics (LEHD) linked employee-employer database, the QWI provides information on key labor market outcomes. In particular, the QWI provides information at the state, industry and national level on the number of hires, separations, job gains and losses as well as average earnings. The QWI defines job gains or job creation at a firm as the non-negative change in employment within a quarter, which can be formally written as:

$$\text{Job Gains } = \max\{Emp_t^{end} - Emp_t^{beginning}, 0\}$$

In contrast, hires at a firm in quarter t is defined at the total number of new employees at a firm that did not have earning in period $t-1$ but that reported earnings at that firm in period t. The measure of hires records the gross inflows into a firm, while the measure of job gains records the non-negative net employment change at the firm. Replacement hires are defined in the QWI as the hires that continue into the next quarter *in excess* of job gains at a firm. Using data from the QWI, we calculate the replacement hiring share as the fraction of hires that are replacement hires, i.e.

$$\text{Replacement Hiring Share} = \frac{\text{Replacement Hires}}{\text{Hires}} = \frac{\text{Hires At End - Job Gains}}{\text{Hires}}$$

where "Hires At End" refer to hires that continue into the next quarter, i.e. an individual who records earnings in periods t and $t+1$ but not in period $t-1$. Importantly, "Hires At End" are a sub-set of "Hires". The latter includes both hires that continue into the next quarter and individuals hired only for that particular quarter, i.e. the individual only has a record of earnings at time t. By definition, when there are zero hires that continue into the next quarter, there would be zero replacement hires recorded.[7] It is important to note that because replacement hires capture the hires that continue into next quarter in excess of job gains, the replacement hiring share is not equivalent to the ratio of separations to hires.

[7]This implies that the level of replacement hires is bounded below by zero. Thus, if a firm contracts and only observed separations, the QWI records zero rather than "negative" replacement hires at this firm.

This is because only a subset of separations are associated with replacement hires.[8] Thus while all replacement hires are associated with separations, not all separations are replacement hires.

Because a replacement hire always coincides with a separation, it is useful to distinguish between the different types of separations that can lead to a replacement hire. Firstly, a replacement hire can occur when the firm conducts on-the-job search and decides to hire a more productive applicant to replace its current incumbent worker. A replacement hire can also occur for reasons unrelated to firm on-the-job search. In particular, a replacement hire can occur whenever the firm re-fills a vacated position. Using the JOLTS micro-data, Elsby et al. (2016) focus on firms who have the same employment level τ periods later and measure the cumulative hires rate (solid blue line) and cumulative quits rate (dashed-red line) at such firms. Importantly, since these firms observe zero net employment change, the cumulative hires is equal to cumulative separations and these cumulative hires represent replacement hires. While quits do affect the amount of replacement hiring, Figure 3 from Elsby et al. (2016) shows that not all replacement hires stem from refilling positions vacated by workers who quit. Rather, Figure 3 reveals that a non-trivial wedge exists between the cumulative hires rate and cumulative quits rates (plus other separations)[9], suggesting that a significant portion of replacement hiring also occurs alongside the event of a layoff.

Our measure of replacement hires strictly follows that provided by the QWI. Elsby et al. (2016) use an alternative measure of replacement hires and define it as the minimum of gross hires and quits at a firm. This measure of replacement hires is not the same as the definition in the QWI but is consistent with the model presented in Elsby et al. (2016) which views replacement hires as being conducted whenever a worker quits. For our purposes, however, the definition of replacement hires as measured in the QWI is more appropriate since it captures replacement hires that are conducted both for the purposes of firm on-the-job search as well as for refilling a vacated position. To see this, consider a firm which had zero workers quit. The firm, however, decided to replace 1 worker with a higher productivity applicant. In this example, there is 1 hire, 1 fire and hence 1 replacement hire. Using a measure of replacement hires which is the minimum of gross hires and quits, however, would suggest that there are zero replacement hires when there are zero quits. As such, the QWI's definition of replacement hires better suits our model's examination of replacement hiring that occurs for both firm on-the-job search reasons as well as for the purpose of refilling vacated positions.

Having defined replacement hires, we use the above definition to construct the replacement hiring

[8] The following accounting identity from the QWI makes clear that replacement hires is not equal to total separations observed in the data:

$$\text{Hires } - \text{Separations} = \text{Job Gains } - \text{Job Losses}$$

Since replacement hires are only measured as the hires in excess of job gains which only counts non-negative net-employment change, replacement hires are not equal to total separations. To see this, consider the example of a firm who started the period with 1 worker. Suppose that worker left the firm and the firm hired a new worker to re-fill its vacated position. In addition, this new worker left the firm before the end of the period. In this example, the firm experienced a net employment change of -1, stemming from the 2 separations and 1 hire. Because that hire did not continue into the next period, the replacement hiring share at this particular firm would be equal to 0 since no hire continued into the next period. However, the ratio of separations to hires in this example would be equal to 2.

[9] The JOLTS data series defines other separations as separations stemming from retirements as well as discharges due to reasons of disability.

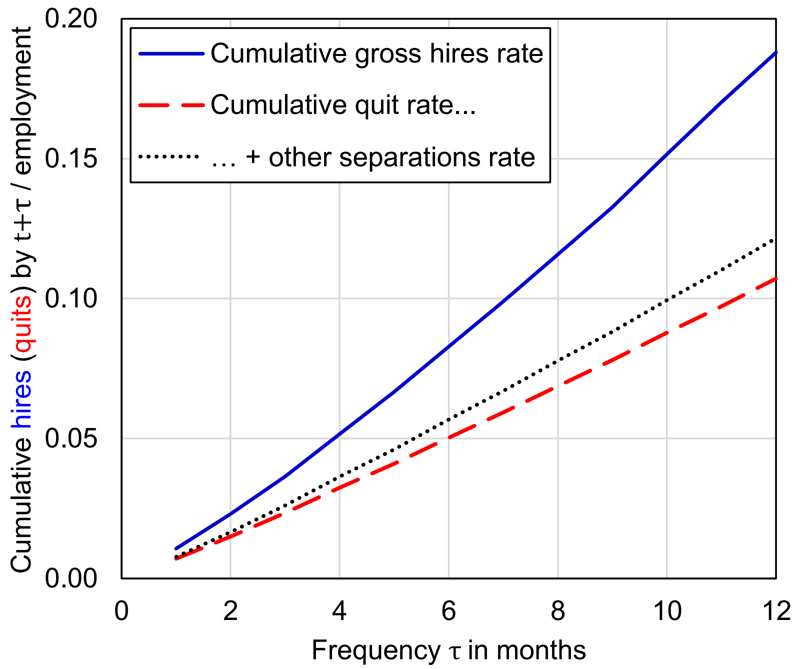


Figure 3: Both Layoffs And Quits Affect Replacement Hiring
Source: Elsby et al. (2016)

share in Figure 2a.[10] Overall we find that the replacement hiring share[11] rose from about 0.33 in the 1990s to about 0.41 in 2017.[12] Separately, Appendix A shows that the rise in the share of replacement hiring is not limited to a single industry's experience but occurs broadly across all industries. Further, Appendix A also shows that the results from a shift-share analysis suggest that the "within" component accounts for the bulk of the increase in the replacement hiring share. This is regardless of whether one cuts the data by industry, firm size or firm age.

These findings indicate that the standard labor search model may be missing an important feature which is on-the-job search by firms. We outline in our model section how replacement hiring can occur

[10] It should be noted that Figure 2a records the trend in the replacement hiring share and not the replacement hiring rate where the denominator in the latter is total employment and the denominator in the former is hires. Further, the replacement hiring share should not be confused with churn in the economy which - following the definition in Burgess et al. (2000)- is defined as total worker flows (hires plus separations) in excess of job flows (hires minus separations). Churn rates at time t are then given by:

$$\text{Churn}_t = \frac{\text{Worker Flows}_t - \text{Job Flows}_t}{0.5(Emp_t + Emp_{t+1})} = \frac{\text{Hires}_t + \text{Separations}_t - (\text{Hires}_t - \text{Separations}_t)}{0.5(Emp_t + Emp_{t+1})} = \frac{2\text{Separations}_t}{0.5(Emp_t + Emp_{t+1})}$$

Since replacement hires are only a subset of all separations, replacement hiring can only be a portion of total churn.

[11] Importantly, Elsby et al. (2016), using their measure of replacement hires as the minimum of gross hires and quits, show in their paper that the replacement hiring share as captured by quits is relatively constant. This is suggestive that the growth in the replacement hiring share may not be driven by increased worker on-the-job search, and instead may be affected by firm on-the-job search.

[12] Our finding that the replacement hiring share rose over time is robust even if we use a different measure of the replacement hiring share. Specifically, if we only consider replacement hires as a fraction of all hires that continue into next quarter (as opposed to all hires that include individuals who were hired in a quarter but who did not stay on until the next quarter), this alternative replacement hiring share rose from 0.55 in the 1990s to about 0.60 in 2017.

in an environment where firms can do on-the-job search and when vacancies do not expire immediately.

3 Model

Workers and firms Time is continuous and runs forever. The economy comprises of a unit mass of infinitely-lived workers who are ex-ante identical. All workers are risk neutral and discount the future at a rate $\rho > 0$. Workers can either be employed or unemployed. Unemployed workers receive flow utility $b \geq 0$ per unit time. The other agents in the economy are firms each of which can employ at most one worker at any date. A firm-worker pair with match quality x produces x units of output at each date. The match quality $x \in [0, \overline{x}]$ is drawn from a time invariant distribution $\Pi(x)$ at the time the firm and worker meet and remains constant for the duration of the match.[13]

Vacancies and Firms Search is random. A firm that decides to enter the market must incur a fixed cost χ to post a vacancy. Posting a vacancy in our model is synonymous with creating a job position. All firms enter the labor market initially as unfilled vacancies. Importantly, unlike the standard DMP setup, in our model, unfilled vacancies do not expire or are destroyed instantly. Instead, vacancies are destroyed at a rate $\delta > 0$. This implies that a vacancy that goes unmatched today can still contact an applicant in the future as long as the vacancy/job position has not been destroyed. In addition, firms who were previously matched to a worker but whose worker separated from them at exogenous rate s, can still transition to become unfilled vacancies so long as their job position has not been destroyed. Thus, firms can replace workers who separated from them without posting a new vacancy.

Because vacancies are synonymous with job positions in our model, firms whose vacancies have been filled and whose job position has not been destroyed, referred to as *recruiting matched firms*, can still continue to meet and accept new applicants, i.e. firms can conduct on-the-job search. If the matched firm chooses to replace its current worker with the new job applicant, it releases its current worker into unemployment. In the case where the worker leaves the firm and the firm is unable to find a replacement, recruiting matched firms become unfilled vacancies. If a firm with an unfilled vacancy hires a worker, it becomes a recruiting matched firm.

Finally, a vacancy or position is destroyed at rate δ. When a unfilled vacancy is exogenously destroyed (at rate δ), the vacancy ceases to exist, while when a currently filled vacancy experiences the same shock, both the existing match and the vacancy cease to exist. This shock can be thought of as a firm no longer needing a worker for a particular position.

Labor Market Both unemployed and employed workers can make contact with vacancies. The rate at which unemployed workers meet vacancies (both currently unfilled and filled) is denoted by p while currently employed workers can conduct on-the-job-search and meet vacancies at a rate $\lambda_w p \leq p$. Similarly, an unfilled vacancy meets job-seekers (both currently unemployed and employed) at a rate q. Unlike in the standard model, firms with currently filled vacancies can also search while on-the-job and meet applicants at a rate $\lambda_f q \leq q$. The job-finding rate p and the job-filling-rate q are determined by a

[13] The support of x is allowed to be unbounded above, i.e., $\overline{x}$ can be ∞. In fact, in our calibrated model, we assume that x is described by a log-normal distribution and hence, the support is unbounded above.

meeting technology which takes as its inputs total vacancies and total applicants:

$$M = v^{1-\alpha}\ell^{\alpha} \tag{1}$$

where ℓ denotes job-seekers and $v = v^u + \lambda_f v^m$ is the total number of vacancies that can be contacted. Here, vacancies include all the unfilled vacancies, v^u, and the fraction of the currently matched firms who receive an opportunity to search, $\lambda_f v^m$. Similarly, $\ell = u + \lambda_w(1-u)$ is the total number of applicants who are available for matching at any date. This includes all the unemployed workers, u, and the fraction of currently employed workers $1-u$ who get a chance to search on-the-job, $\lambda_w(1-u)$.[14]

Importantly, meeting rates are not equivalent to hiring rates. In order for a meeting to result in a hire, both the firm and the worker must agree to form a match. If the match is between an unfilled vacancy and an unemployed worker, then both parties agree to form a match as long as the match-specific productivity drawn is above a threshold $\widetilde{x}$, which is determined in equilibrium.[15] However, if the unemployed individual meets a recruiting matched firm, the new match quality x drawn must be at least as large as the incumbent worker's match quality. Thus, although the rate with which an unemployed worker meets a filled and unfilled vacancy is the same, the probability with which she will be hired is (weakly) lower for vacancies which are already filled. Similarly, if an unfilled vacancy meets a currently employed worker, the new match is only formed if the new match productivity drawn exceeds that of the employed worker's old match. Finally, a meeting between a currently employed worker and a recruiting matched firm, i.e. a filled vacancy, results in a new match only if the new match-productivity drawn exceeds that observed in both existing matches.

Workers can be both exogenously and endogenously separated from firms. The former occurs at rate s, while the latter occurs whenever a recruiting matched firm replaces their current worker with a better applicant. Similarly, a filled vacancy can be both exogenously and endogenously separated from their employee. Exogenous separations occur at rate s, while endogenous separations occur when the worker succeeds in finding a better match while searching on-the-job. Having described the environment, we next describe the firms' problems.

3.1 Firm's Problem

Recruiting matched firms The value of a recruiting matched firm with current match quality x can be written as:

$$\rho J(x) = x - w(x) + \delta\left[J^0 - J(x)\right] + \Big(s + p^*(x)\Big)\left[J^u - J(x)\right] + R(x) \tag{2}$$

The firm receives current profits $x - w(x)$ and can undergo four possible events in the future. First, the vacancy/job position is destroyed at a rate δ, causing the match and vacancy to be destroyed. The firm now has the choice of creating a new unfilled vacancy and hence the change of value is $J^0 - J(x)$ where J^0 denotes the value of creating a new vacancy. Second, it may undergo an exogenous separation at rate s and become an unfilled vacancy with the associated change in value $J^u - J(x)$. Third, the firm's current worker successfully searches on-the job and leaves the current firm to join another match.

[14] Of course, consistency requires that $1-u = v^m$.

[15] See Section 3.5 for details.

As in the previous case, the firm transitions into an unfilled vacancy when this event occurs. $p^*(x)$ denotes the rate at which a worker with current match-productivity x successfully finds a better job while searching on-the-job and is formally given by:

$$p^*(x) = \lambda_w p \left\{ \left(\frac{v^u}{v}\right) [1 - \Pi(x)] + \left(\frac{\lambda_f v^m}{v}\right) \left[[1 - \Pi(x)] F(x) + \int_x^{\overline{x}} [1 - \Pi(\varepsilon)] \, dF(\varepsilon) \right] \right\} \quad (3)$$

The events that lead a worker to endogenously separate from the firm can be described as follows: first, a currently employed worker can meet a vacancy at rate $\lambda_w p$. Conditional on a meeting, the worker meets an unfilled vacancy with probability v^u/v and forms a match as long as the new match quality is higher than its current x (which occurs with probability $1-\Pi(x)$). With probability $1-v^u/v = \lambda_f v^m/v$, the worker meets a filled vacancy with match quality ε; with probability $F(x)$ the filled vacancy's current match quality is lower than x and the new match is only formed if the pair draw a new match-quality larger than the workers current match-quality x (which occurs with probability $1 - \Pi(x)$). Here, $F(\cdot)$ denotes the endogenous cumulative distribution function of existing matched firm-worker pairs across match quality with $F(\widetilde{x}) = 0$ and $F(\overline{x}) = 1$.[16] The last term $\int_x^{\overline{x}} [1 - \Pi(\varepsilon)] \, dF(\varepsilon)$ represents the probability that a new match is formed when the employed worker with current match-quality x meets a filled vacancy with match quality $\varepsilon > x$ in which case the match is only formed if the new match-quality is larger than ε.

Returning to equation 2, the last term in that equation, $R(x)$, denotes the expected value of on-the-job search by a firm with current match quality x:

$$\begin{aligned} R(x) = \; & \lambda_f q \Big[\left(\frac{u}{\ell}\right) \int_x^{\overline{x}} [J(y) - J(x)] \, d\Pi(y) \\ & + \left(\frac{\lambda_w v^m}{\ell}\right) \left\{ \int_x^{\overline{x}} [J(y) - J(x)] \, d\Pi(y) \, F(x) + \int_x^{\overline{x}} \int_\varepsilon^{\overline{x}} [J(y) - J(x)] \, d\Pi(y) \, dF(\varepsilon) \right\} \Big] \end{aligned} \quad (4)$$

A recruiting matched firm has effective rate $\lambda_f q$ of conducting on-the-job search. Conditional on the meeting, the first term is the change in value associated with the event when the firm meets an unemployed worker, draws a new match-quality $y > x$, forms the new match and enjoys a change of value $J(y) - J(x)$. The term on the second line reflects the expected change in value when the recruiting matched firm meets another already employed worker and a new match is formed. The first-term on the second line reflects the event when the firm with current match-quality x meets a currently employed worker who has a match quality $\varepsilon < x$ with her incumbent firm (this happens with probability $F(x)$). In this case, the worker currently with match-quality ε is always willing to form the new match if the firm with match quality $x > \varepsilon$ is willing to do so. Similarly, the second term on the second line refers to the event whenever the firm witn match quality x meets an employed worker with match-quality $\varepsilon \geq x$. In this case, its the worker's decision to form a match which is binding.

[16] We describe $F(\cdot)$ in more detail in the subsequent sections.

Firms with unfilled vacancies The value of an unfilled vacancy can be written as:

$$\rho J^{u}=\delta\left[J^{0}-J^{u}\right]+q\left\{\left(\frac{u}{\ell}\right)\int_{\widetilde{x}}^{\overline{x}}\left[J\left(y\right)-J^{u}\right]d\Pi\left(y\right)+\left(\frac{\lambda_{w}v^{m}}{\ell}\right)\int_{\widetilde{x}}^{\overline{x}}\int_{\varepsilon}^{\overline{x}}\left[J\left(y\right)-J^{u}\right]d\Pi\left(y\right)dF\left(\varepsilon\right)\right\} \quad (5)$$

In the current period, an unfilled vacancy observes zero production and zero current profits. The vacancy is destroyed at rate δ in which case it ceases to exist (associated with the change of value $J^0 - J^u$). At rate q, the firm meets an applicant and with probability $\frac{u}{\ell}$, this applicant is currently unemployed. The unemployed worker and unfilled vacancy form a match as long as the match-quality, y, drawn is above the reservation match quality denoted by $\widetilde{x}$. The firm's gain from such a match is given by $J(y) - J^u$. The reservation match-productivity $\widetilde{x}$ is defined as the lowest value of x for which the value of the match is non-negative:

$$J(\widetilde{x}) - J^u = 0 \quad (6)$$

In the complementary case, i.e. with probability $1 - u/\ell = \lambda_w v^m/\ell$ the unfilled vacancy meets an employed worker with current match-quality ε. They form a new match only if the they draw a new match-quality $y > \varepsilon$ in which case the firm's change in value is $J(y) - J^u$.

Free Entry We assume that there is free-entry, i.e. the value of creating a new vacancy or job position is driven down to 0:

$$J^0 = -\chi + J^u = 0 \quad (7)$$

which implies that the value of an unfilled vacancy $J^u = \chi$ is positive. In other words, an unexpired vacancy provides the firm with a positive option value and affords the firm the ability to continue to search tomorrow even if it rejects or fails to meet a worker today. Further - and as we discuss subsequently - this positive option value allows raises the recruiting matched firm's outside option when bargaining with the worker, allowing it to bargain wages down.

3.2 Worker's Problem

Unemployed workers The asset value of a worker from unemployment, U, can be written as:

$$\rho U \;=\; b+p\left\{\left(\frac{v^{u}}{v}\right)\int_{\widetilde{x}}^{\overline{x}}\left[W\left(y\right)-U\right]d\Pi\left(y\right)+\left(\frac{\lambda_{f}v^{m}}{v}\right)\int_{\widetilde{x}}^{\overline{x}}\int_{\varepsilon}^{\overline{x}}\left[W\left(y\right)-U\right]d\Pi\left(y\right)dF\left(\varepsilon\right)\right\} \quad (8)$$

where $F(\varepsilon)$ denotes the fraction of recruiting matched firms whose worker possesses match quality ε or lower. $W(y)$ is the value of being employed at a recruiting firm with match quality y.

The value of unemployment can be decomposed into two terms: b, the flow utility associated with home production and the second term in equation (8) which denotes the expected change in value that the worker enjoys in the event that he transitions to employment in the future. At a rate p, an unemployed worker meets a vacancy. With probability v^u/v, this vacancy is currently unfilled and the worker is accepted whenever he draws a match quality higher than $\widetilde{x}$. However, with probability

$\lambda_f v^m/v$, the unemployed worker encounters a recruiting matched firm and is only hired when she draws a match quality y that is higher than the incumbent's value. The second term inside the parenthesis captures the unemployed worker's change in value when she is accepted by a recruiting matched firm with current match quality ε weighted by the probability of meeting such a firm.

Unlike the standard model, it is useful to note that the introduction of firm on-the-job search affects the composition of vacancies which in turn affects the worker's value of unemployment. When $\lambda_f v^m/v$ is high, unemployed workers are more likely to encounter recruiting matched firms as opposed to unfilled vacancies. Because recruiting matched firms require unemployed applicants to draw a match productivity above their incumbent worker's match quality, the wedge between meeting and hiring rates is larger when the composition of vacancies tilts towards recruiting matched firms. As such, a higher $\lambda_f v^m/v$ proxies for a decline in *measured matching efficiency* and acts towards reducing the worker's value of unemployment.

Employed workers The asset value of an employed worker at a recruiting firm with match quality x can be written as:

$$\rho W(x) = w(x) - \Big(\delta + s + q^*(x)\Big)[W(x) - U] + H(x) \tag{9}$$

where $w(x)$ denotes the wages paid. There are three events that transition the worker into unemployment and hence result in a change in value of $-[W(x) - U]$. First, at rate δ, the vacancy/position is destroyed, and the worker transitions to unemployment. Second, the worker is exogenously displaced into unemployment at rate s. Finally, the worker is also endogenously separated into unemployment when its firm meets a new applicant and forms a new match with probability $q^*(x)$ which is formally given by:

$$q^*(x) = \lambda_f q\left\{\left(\frac{u}{\ell}\right)[1-\Pi(x)] + \left(\frac{\lambda_w v^m}{\ell}\right)\left[[1-\Pi(x)]F(x) + \int_x^{\overline{x}}[1-\Pi(\varepsilon)]\,dF(\varepsilon)\right]\right\} \tag{10}$$

The events which up to the worker being endogenously displaced into unemployment are as follows: its matched firm meets an applicant at rate $\lambda_f q$. Conditional on a meeting, the firm meets an unemployed applicant with probability u/ℓ and forms a match as long as the new match quality is higher than its current x (which occurs with probability $1-\Pi(x)$). With probability $1 - u/\ell = \lambda_w v^m/\ell$ the firm meets a currently employed worker with match quality ε; with probability $F(x)$ the worker's current match quality, ε, is lower than x and a new match is only formed if the pair draw a new match-quality, y, larger than the firm's current x (which occurs with probability $1-\Pi(x)$). The last term $\int_x^{\overline{x}}[1-\Pi(\epsilon)]\,dF(\epsilon)$ represents the probability that a new match is formed when the firm with match-quality x meets a currently employed worker with match quality $\varepsilon > x$ in which case the match is only formed if the new match-quality is larger than ε.

In addition, the value of an employed worker with match-quality x also includes the value of worker on-the-job search as denoted by $H(x)$ where:

$$\begin{aligned} H\left(x\right) &= \lambda_w p \Big\{ \left(\frac{v^u}{v}\right) \int_x^{\overline{x}} \left[W\left(y\right)-W\left(x\right)\right] d\Pi\left(y\right) \\ &+ \left(\frac{\lambda_f v^m}{v}\right) \left[\int_x^{\overline{x}} \left[W\left(y\right)-W\left(x\right)\right] d\Pi\left(y\right) F\left(x\right) + \int_x^{\overline{x}} \int_\varepsilon^{\overline{x}} \left[W\left(y\right)-W\left(x\right)\right] d\Pi\left(y\right) dF\left(\varepsilon\right) \right] \Big\} \end{aligned} \tag{11}$$

A worker searching on the job meets a vacancy at rate $\lambda_w p$. Conditional on meeting, the first term is the change in value experienced by the worker with current match-quality x when she meets an unfilled vacancy, draws a new match-quality $y > x$ and forms a new match. This results in a change of value of $W(y) - W(x)$. The term on the second line reflects the change in value when the currently employed workerwith match-quality x meets an a recruiting matched firm and forms a new match. The first-term on the second line reflects the event when the worker meets a currently filled vacancy who has a match quality $\varepsilon < x$ (this happens with probability $F(x)$). They form a match as long as the new match-quality is greater than x. Similarly, the second term on the second line refers to the event whenever the employed worker meets a filled vacancy with match-quality $\varepsilon \geq x$. In this case, the new match is only formed if the pair draw a match-quality higher than ε.

Unlike the standard model, the introduction of firm on-the-job search introduces additional *job insecurity* for the worker through endogenous firm separations. In this case, holding all else constant, if the ease of firm on-the-job search increases, i.e. λ_f rises, then $q^*(x)$ rises and workers' employment spells are shortened. This has the effect of lowering workers' employment values and feedbacks into lowering workers' outside options, ρU, through the expected change in value when unemployed workers form a match.

3.3 Surplus and Wage Formation

The joint payoff to a match is the total surplus shared by the firm and worker relative to continuing to search: $S(x) = J(x) - J^u + W(x) - U$. Manipulating equations (2),(5),(8) and (9), one can show that the surplus for a matched firm-worker pair with match quality x is given by:

$$\Big(\rho + s + \delta + \eta q^*(x) + (1-\eta) p^*(x)\Big) S(x) = x - \rho U + R\left(x\right) + H\left(x\right) - \left(\rho + \delta\right)\chi \tag{12}$$

where $\rho U, q^*(x)$ and $p^*(x)$ are as defined in (8),(10) and (3) respectively. The first thing to notice in the equation above is that the higher the outside option of firms (higher χ)[17], the smaller is the surplus because a higher outside option lowers a firm's gain from matching since it can always afford to wait and meet a worker with higher match-quality. Similarly, a higher value of unemployment, ρU, given everything else, lowers the surplus since the worker's gain to matching is lower. (12) also shows that holding all else constant, the surplus from a match is increasing in the firm's and the worker's expected value, $R(x)$ and $H(x)$ respectively, from searching on-the job.

[17]recall that in equilibrium, the value of a filled vacancy is given by $J^u = \chi$.

Wage Determination Wages are determined at each date via Nash Bargaining:

$$w(x) = \underset{w(x)}{\arg\max} \Big[J(x) - J^u\Big]^{1-\eta} \Big[W(x) - U\Big]^{\eta} \tag{13}$$

where $\eta \in [0, 1]$ denotes the bargaining power of a worker. Bargaining over wages takes place only after matches have been formed. This implies that whenever a recruiting matched firm chooses to hire a new applicant, he releases his current worker into unemployment prior to bargaining with the new applicant. Similarly, whenever a currently employed worker chooses to form a match with a different vacancy, she first vacates her current job. We further assume that there are no recalls. As such, when the firm with match quality x and a new applicant bargain over wages, the firm's outside option is simply the positive option value of an unfilled vacancy, J^u, and not $J(x)$. Similarly, for an employed individual with current match quality y, the worker's outside option is U and not $W(y)$.[18] As is well known, the Nash bargaining solution implies that the surplus is split between firm and worker such that:

$$J(x) - \chi = (1 - \eta)S(x) \qquad \text{and} \qquad W(x) - U = \eta S(x) \tag{14}$$

3.4 Labor Market Flows

Having described the relevant value functions, we proceed to describe labor market flows next.

Unemployed The steady state rate of unemployment u satisfies:

$$q\left(\frac{u}{\ell}\right)\left[1 - \Pi\left(\widetilde{x}\right)\right] v^u = (s + \delta)\, v^m + 2q\left(\frac{\lambda_f v^m \times \lambda_w v^m}{\ell}\right) \int_{\widetilde{x}}^{\overline{x}} \left[1 - \Pi\left(z\right)\right] F\left(z\right) dF\left(z\right) \tag{15}$$

The LHS represents the outflows from the pool of unemployed. At rate qu/ℓ, an unfilled vacancy meets an unemployed worker and hires her if they draw a match quality above $\widetilde{x}$. Notice that there is no net outflow when a currently filled vacancy hires an unemployed worker, as it also releases its current worker into unemployment. This feature of firm on-the-job search distinguishes it from worker on-the-job search. Note that when workers conduct on-the-job search, they leave their current firm, causing an unfilled vacancy to open up and creating the start of a vacancy chain.[19] Here, if a firm conducts on-the-job search and hires a worker out of unemployment, it does not create a vacancy chain nor does it affect the unemployment pool on net, since hiring a worker out of unemployment requires it to displace its incumbent worker into unemployment.

The RHS denotes the flows into unemployment. The first term on the RHS, $(s + \delta)v^m$, is the fraction of all currently employed workers who experience a exogenous separation s or who observe

[18] Notice that this assumption is without loss of generality since wages are determined via Nash Bargaining each period without commitment. Even if firms bargained before separating with their current worker, and effectively used $J(y)$ as their outside option, it would mean that at the instant the next match is formed, it would revert to having J^u as its current outside option and would have to pay workers wages commensurate with equation (13). Of course, if the the environment featured commitment by agents in the form of long-term contracts, then this would not be true and in that scenario, currently matched recruiting firms would offer a lower wage than unfilled vacancies for an applicant with the same match quality.

[19] Both Elsby et al. (2016) and Mercan and Schoefer (2019) explore how worker on-the-job search can give rise to vacancy chains, which are defined as the phenomenon where vacancies beget more vacancies.

their position/vacancy being destroyed (δ). The second term on the RHS refers to the flows into unemployment when a currently employed worker forms a match with a currently filled vacancy. In this case, the employed worker displaces the filled vacancy's incumbent worker into unemployment.

Unfilled vacancies Since vacancies are long-lived, the stock of unfilled and unexpired vacancies, v^u in steady state is implicitly defined by:

$$q\left(\frac{u}{\ell}\right)\left[1-\Pi\left(\widetilde{x}\right)\right]v^u+\delta v^u=v^{new}+sv^m+2q\left(\frac{\lambda_f v^m\times\lambda_w v^m}{\ell}\right)\int_{\widetilde{x}}^{\overline{x}}\left[1-\Pi\left(z\right)\right]F\left(z\right)dF\left(z\right) \tag{16}$$

The LHS of (16) represents the outflow from the pool of unfilled vacancies. The first term on the LHS $q\left(\frac{u}{\ell}\right)\left[1-\Pi\left(\widetilde{x}\right)\right]v^u$ is the number of unfilled vacancies which met an unemployed worker and formed a match. Notice that when an unfilled vacancy poaches a currently employed worker, there is no net-outflow from the pool of unfilled vacancies since the worker leaves the pre-existing match, transitioning that vacancy into an unfilled vacancy. The second term on the LHS δv^u is the number of unfilled vacancies which are destroyed.

The RHS of (16) represents the inflow into the pool of unfilled vacancies. The first component of inflows is v^{new}, the flow of newly created vacancies. Importantly, v^{new} is not counted as part of the vacancies available for matching today.[20] In the continuous time limit, $\theta=(v^u+\lambda_f v^m)/\ell\equiv v/\ell$. Thus, workers can only match with existing/old vacancies. New vacancies only add to the stock of unfilled vacancies in the future. The second term on the RHS sv^m denotes all matched vacancies which experience an exogenous separation with their current worker. Finally, the third term on the RHS represents the flows into the pool of unfilled vacancies when a currently matched vacancy and a currently employed worker form a new match, displacing the old vacancy which employed the worker into unfilled status.

[20]The total vacancies available for matching at time t are given by $v_t=(1-\delta\Delta)(v^u_{t-\Delta}+\lambda_f v^m_{t-\Delta})+v^{new}_t\Delta$ where Δ is the length of one period, $(1-\delta\Delta)(v^u_{t-\Delta}+\lambda_f v^m_{t-\Delta})$ is the stock of unexpired vacancies from the end of period $t-\Delta$. v^{new}_t is the number of new vacancies posted per unit time. Since each period is Δ units long, the total number of new vacancies posted in period t is $v^{new}_t\Delta$. Thus, market tightness can be written as:

$$\theta_t=\frac{(1-\delta\Delta)(v^u_{t-\Delta}+\lambda_f v^m_{t-\Delta})+v^{new}_t\Delta}{u_{t-\Delta}+\lambda_w v^m_{t-\Delta}}$$

In the continuous time limit, $\Delta\to 0$, the term $v^{new}\Delta$ becomes vanishingly small implying that current vacancies available for matching at period t consist only of existing or *old* vacancies $v^u_t+\lambda_f v^m_t$.

Endogenous distribution of match-productivity The steady state distribution of matched firm-worker pairs across match qualities $F(x)$ is implicitly given by:

$$
\begin{aligned}
q\left(\frac{u}{\ell}\right)\left[\Pi\left(x\right)-\Pi\left(\widetilde{x}\right)\right]v^{u} \quad = \quad & \left(s+\delta\right)F\left(x\right)v^{m}+q\left(\frac{u}{\ell}\right)F\left(x\right)\lambda_{f}v^{m}\left[1-\Pi\left(x\right)\right]+p\left(\frac{v^{u}}{v}\right)F\left(x\right)\lambda_{w}v^{m}\left[1-\Pi\left(x\right)\right] \\
& +2q\left(\frac{\lambda_{f}v^{m}\lambda_{w}v^{m}}{\ell}\right)\left\{\left[1-\Pi\left(x\right)\right]F\left(x\right)+\int_{x}^{\overline{x}}\left[1-\Pi\left(\varepsilon\right)\right]dF\left(\varepsilon\right)\right\}F\left(x\right) \\
& +q\left(\frac{\lambda_{f}v^{m}\lambda_{w}v^{m}}{\ell}\right)\left\{\int_{\widetilde{x}}^{x}\int_{z}^{x}\left[\Pi\left(x\right)-\Pi\left(\varepsilon\right)\right]dF\left(\varepsilon\right)dF\left(z\right)\right. \\
& \left.+\int_{\widetilde{x}}^{x}\left[\Pi\left(x\right)-\Pi\left(z\right)\right]F\left(z\right)dF\left(z\right)\right\} \qquad \text{for } x\in\left(\widetilde{x},\overline{x}\right)
\end{aligned}
\tag{17}
$$

with $F(\widetilde{x}) = 0$ and $F(\overline{x}) = 1$. The LHS of (17) represents the inflow into the set of matched-vacancies with match quality between $\widetilde{x}$ and x and is the number of unfilled vacancies which match with an unfilled worker and draw a match quality above $y \in [\widetilde{x}, x]$.

The RHS of (17) denotes the outflows from the same set. The first term on the RHS denotes the number of matched-vacancies with match-quality less than x, i.e. $(F(x)v^m)$, who experience an exogenous separation or destruction of the vacancy. The second term represents the number of currently matched-vacancies who successfully matched with an unemployed worker and drew a new match-quality above x, thus reducing the number of matches with match-quality below x. The third term is the number of currently employed workers with match-quality less than x who successfully match with an unfilled vacancy and draw a new match-quality greater than x. The next two lines of (17), similarly, describe various events where a matched-vacancy and employed worker form a match and either result in one or two matches being moved out of the set of vacancies with match quality less than x. The third (and fourth) line of (17) describes the case where an employed worker with $\epsilon < x$ and a filled vacancy with $z < x$ meet, and they draw $\max\{z, \epsilon\} < y \leq x$, then one firm-worker pair leaves the distribution of firm-worker pairs with match quality less than or equals to x. The second line of (17) describes the case where an employed worker with $\epsilon < x$ and a filled vacancy with $z < x$ meet, and they draw $y > x$, then two firm-worker pairs leave the measure of matched pairs with match quality less than or equals to x.

The distribution of matches by quality $F(x)$ is informative about the replacement hiring share. If the distribution of matches are skewed towards low values of x, then there is substantial room for matched-firms to find a better match and thus to conduct replacement hiring. Similarly, employed workers also have substantial room to find a better match by searching on-the-job. When these workers find better matches and leave the firm, the firm with an unfilled vacancy must find a replacement, again encouraging replacement hiring. In contrast, if the distribution of matches is concentrated around higher values of x, both matched firms and employed workers find it harder to find better matches, reducing replacement hiring.

3.5 Closing the Model

The entire model so far has been summarized by the surplus equations and the labor market flows. However, all these relationships depend critically on the reservation match quality, $\widetilde{x}$, and the job-filling rate q which in turn is a function of labor market tightness θ. Lemma 1 summarizes the key equations which pin down the equilibrium $(\tilde{x}, \theta)$:

Lemma 1. *In steady state, the equilibrium $\tilde{x}$ and θ are determined by the following equations:*

$$(\rho+\delta)\chi = (1-\eta) q \left[\left(\frac{u}{\ell}\right) \int_{\widetilde{x}}^{\overline{x}} S(y)\, d\Pi(y) + \left(\frac{\lambda_w v^m}{\ell}\right) \int_{\widetilde{x}}^{\overline{x}} \int_{\varepsilon}^{\overline{x}} S(y)\, d\Pi(y)\, dF(\varepsilon) \right] \tag{18}$$

$$\begin{aligned} \widetilde{x} &= \rho U - \eta \lambda_w p \left[\left(\frac{v^u}{v}\right) \int_{\widetilde{x}}^{\overline{x}} S(y)\, \pi(y)\, dy + \left(\frac{\lambda_f v^m}{v}\right) \int_{\widetilde{x}}^{\overline{x}} S(y)\, F(y)\, \pi(y)\, dy \right] \\ &\quad + (1-\lambda_f) q (1-\eta) \left[\left(\frac{u}{\ell}\right) \int_{\widetilde{x}}^{\overline{x}} S(y)\, \pi(y)\, dy + \left(\frac{\lambda_w v^m}{\ell}\right) \int_{\widetilde{x}}^{\overline{x}} S(y)\, F(y)\, \pi(y)\, dy \right] \end{aligned} \tag{19}$$

where $S(x)$ and $F(x)$ are implicitly defined in equations (12) and (17). The asset value of unemployment ρU is given by:

$$\rho U = b + \eta p \left\{ \left(\frac{v^u}{v}\right) \int_{\widetilde{x}}^{\overline{x}} S(y)\, d\Pi(y) + \frac{\lambda_f v^m}{v} \int_{\widetilde{x}}^{\overline{x}} \int_{\varepsilon}^{\overline{x}} S(y)\, d\Pi(y)\, dF(\varepsilon) \right\}$$

Proof. See Appendix B. □

Equation (18) is the free-entry condition where we have used $J^u = \chi$ from (7) and the solution to the Nash bargaining problem (14): $J(x) - \chi = (1-\eta)S(x)$ in (5). Equation (18) describes the minimum level of match-productivity for which a firm with an unfilled vacancy is willing to form a match for a given θ - or how *selective* a firm is as a function of labor market tightness. Appendix **??** shows that (18) implies a negative relationship between $\widetilde{x}$ and θ - firms are more selective when the rate of contacting applicants is high. In a slack labor market (lower θ, high q), holding out for a better worker is relatively costless for the firm, and hence the firm raises the minimum level of match quality $\widetilde{x}$ for which it is willing to accept a worker. Conversely, in a tight labor market (low q), holding out for a better applicant is more costly as the firm is unlikely to meet another applicant soon. As such, tight labor markets are associated with lower firm selectivity. Figure 4 depicts the firm's free entry curve and the negative relationship between θ and $\widetilde{x}$.

Equation (19) can be thought of as the *worker's indifference condition*: given a level of market tightness θ, (19) defines the reservation match-quality for which a worker will be filling to exit unemployment and form a match.[21] A higher the value of unemployment ρU, the more *selective* a worker is, i.e. a higher $\widetilde{x}$. Since a tighter labor market (higher θ) implies a higher value of unemployment ρU, (19) implies a positive relationship between $\widetilde{x}$ and θ.[22] Greater opportunity for the worker to search on-the-job (higher λ_w) makes workers less selective, because even if the worker accepts a low match-quality and hence low paying job, she can easily search on-the-job and find a better job. Finally, a higher λ_f, or greater ease with which a firm can search on-the-job lowers the worker's reservation $\widetilde{x}$. This is because a higher λ_f allows a firm to replace a worker easily, thus pushing down the wage that a worker receives for any given match-quality x as we show next. Figure 4 depicts the worker's indifference curve and shows the upward sloping relationship between θ and $\widetilde{x}$.

The unique equilibrium level of selectivity $\widetilde{x}$ and labor market tightness θ is given by the intersection

[21](19) is derived by evaluating (12) at $x = \widetilde{x}$ and using $S(\widetilde{x}) = 0$.

[22]See Appendix **??** for a proof.

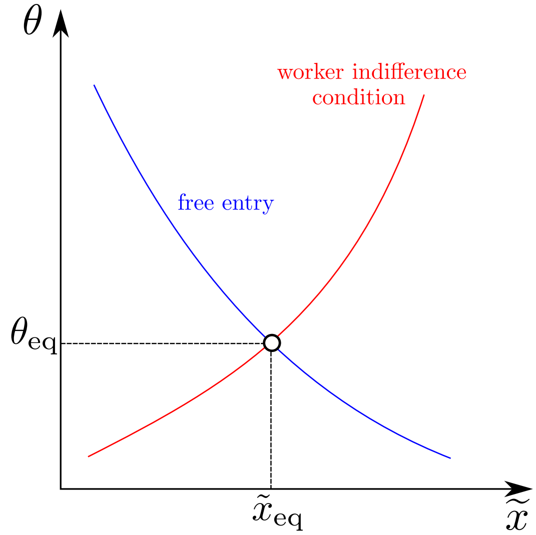


Figure 4: Equilibrium $\widetilde{x}$ and θ

of the two aforementioned curves describing workers' selectivity vs firms' selectivity respectively. Next, we describe how changes in key parameters affect equilibrium outcomes.

4 Forces at Play

Varying the ease with which firms and workers can search on-the-job (λ_f and λ_w resp.) has implications for the behavior of average wages and labor productivity. To identify exactly how a change in the ease of searching on-he-job by workers and firms affects average wages and labor productivity differently, it is useful to work with the two polar cases: where either only worker on-the-job search is operative ($\lambda_f = 0$) or firm on-the-job search is operative ($\lambda_w = 0$). It is useful to consider these special cases, because the model simplifies and admits many closed form expressions under these polar cases allowing us to much better understand the results from our quantitative exercise in Section 5 which works with the general model in which both $\lambda_w, \lambda_f \neq 0$

In what follows, we conduct comparative static exercises where we hold $\widetilde{x}$ and θ constant. We do this because changing either λ_w or λ_f causes both the free entry curve and the worker's indifference condition to shift and knowing what happens to equilibrium θ and $\widetilde{x}$ depends on the relative magnitude to which each curve moves. As such, the following represents a *partial* equilibrium analysis.

4.1 Worker On-the-Job Search Only

We start by shutting off firm on-the-job search, i.e. $\lambda_f = 0$, and consider how variations in λ_w affect the productivity distribution as well as wages. In this special case, the distribution of match-quality among employed firm-worker pairs (given generally by (17)) simplifies and admits a closed form:

$$F(x) = \frac{s+\delta}{(s+\delta+\lambda_w p[1-\Pi(x)])} \frac{\Pi(x) - \Pi(\widetilde{x})}{1-\Pi(\widetilde{x})} \tag{20}$$

A quick inspection of (20) reveals that for a given $(\widetilde{x}, \theta)$, $F(x \mid \lambda^{high})$ first-order stochastically dominates (FOSD) $F(x \mid \lambda^{low})$ for $\lambda^{high} < \lambda^{low}$. Since $F(x \mid \lambda^{high})$ FOSD $F(x \mid \lambda^{low})$, it is clear that the average productivity of all matches is higher with a higher λ_w:[23]

$$\int_{\widetilde{x}}^{\overline{x}} x dF(x \mid \lambda_w^{high}) dx \geq \int_{\widetilde{x}}^{\overline{x}} x dF(x \mid \lambda_w^{low}) dx$$

Intuitively, when workers have greater opportunity (higher λ_w) to conduct on-the-job search, they find it easier to move to jobs higher match-quality jobs which pay higher wages. Consequently, more matches tend to have higher x's. Overall, greater opportunity for workers allows them to easily move to higher x jobs and thus, raises average productivity.

What effect does a change in λ_w have on wages? Given $(\widetilde{x}, \theta)$, a higher λ_w raises the worker's value of search $H(x)$ in (11) and consequently raises the surplus from a match $S(x)$ given by (12). In the polar case with $\lambda_f = 0$, the ODE in (12) can be explicitly solved:[24]

$$S(x) = \frac{1}{\left[\rho + s + \delta + \lambda_w p \left[1 - \Pi\left(x\right)\right]\right]^{1-\eta}} \int_{\widetilde{x}}^{x} \frac{1}{\left[\rho + s + \delta + \lambda_w p \left[1 - \Pi\left(y\right)\right]\right]^{\eta}} dy$$

and the wage paid to a worker in a job with match-quality x is given by:

$$w\left(x\right) = \eta x + \left(1 - \eta\right) \widetilde{x} - \left(\rho + \delta\right) \chi + \left(1 - \eta\right) \eta \lambda_w p \int_{\widetilde{x}}^{x} S\left(y\right) \pi\left(y\right) dy \tag{21}$$

Equation 32 shows that the above wage is increasing in λ_w.[25] Intuitively, a rise in λ_w improves both worker's employment values and their outside options, allowing them to take home a larger wage. Furthermore, the derivative of $w(x)$ w.r.t. x is given by:

$$w'\left(x\right) = \eta + \left(1 - \eta\right) \eta \lambda_w p S\left(x\right) \pi\left(x\right) \tag{22}$$

Equation 22 informs us that the slope of the wage function is increasing in λ_w, implying that wages are rising faster in x when λ_w is higher.[26] Since higher worker on-the-job search leads to more mass at higher productivity levels and since wages are growing even faster in x when λ_w is high, a rise in λ_w works against a widening of the productivity-wage gap. With a higher λ_w, firms have to pass off a larger portion of the greater surplus from a higher x to workers as higher wages. As a result, the equilibrium wage distribution for high λ_w also FOSD's the wage distribution when λ_w is low. In summary, a greater ease for workers to search on-the-job encourages wages to grow at a faster rate with productivity improvements.

[23] This follows from the properties of first order stochastic dominance.

[24] See Appendix B.3 for details.

[25] Technically, surplus is also affected by λ_w. Appendix B.3 shows that the wage function is still increasing in λ_w when we substitute in for surplus.

[26] See Appendix B.3 for a formal proof of the slope of the wage function increasing in λ_w.

4.2 Firm On-the-job Search Only

Next, we consider the opposite polar case, $\lambda_w = 0$, and explore how changing λ_f affects productivity and wages. As in the previous special case, the distribution of match-quality amongst matched firm-worker pairs admits a closed form solution in this special case with $\lambda_w = 0$ but $\lambda_f > 0$. Equation (17) simplifies to yield an expression for $F(x)$:

$$F(x) = \left(\frac{s+\delta}{s+\delta+\lambda_f q\left[1-\Pi(x)\right]} \right) \frac{\Pi(x) - \Pi(\widetilde{x})}{1-\Pi(\widetilde{x})} \tag{23}$$

Equation (23) shows that for a given $(\widetilde{x}, \theta)$, a greater ability of firms to conduct on-the-job (higher λ_f) results in more matches with a higher level of match-quality x: $F(x \mid \lambda_f^{high}) \leq F(x \mid \lambda_f^{low})$ In other words $F(x \mid \lambda_f^{high})$ FOSD's $F(x \mid \lambda_f^{low})$ and so the average productivity is increasing in λ_f. Thus, much like a greater ability for workers to search on-the-job, a greater ability for firms to search on-the-job also tends to increase average productivity. Intuitively, a higher λ_f implies that firms can re-match and form more productive matches more easily and as such, climb the productivity ladder.

However, the effect of a higher λ_f on wages is the opposite. A higher λ_f increases the firm's values of being matched through higher value of search $R(x)$, given by equation (4). This also raises the surplus from a match for any given level x. But a higher λ_f also tends to make ρU and hence the outside option of workers lower. TO see this, notice that a higher λ_f raises job insecurity for the worker as it makes it more likely that the firm will replace her with a more productive worker. Moreover, holding $(\widetilde{x}, \theta)$ constant, a higher λ_f means that a unemployed worker is more likely to meet a filled vacancy rather than an unfilled vacancy. Since she must draw a match-quality higher than that in the firm's incumbent match, this lowers the probability of forming a new match and thus of exiting unemployment. All these forces work towards lowering the value of unemployment and hence the workers outside option. Thus, even though a higher λ_f tends to increase surplus $S(x)$, it tends to reduce the workers outside option.

The above opposing effects - larger surplus but lower worker outside options with higher λ_f- lead to contrasting implications for the level vs. slope of the wage function. On the one hand, equation 24 highlights that the wage level is increasing in λ_f,[27] because increased firm on-the-job search raises surplus. On the other hand, equation 25 makes clear the slope of the wage function is strictly decreasing in λ_f precisely because workers' outside options are worsened when firms have a higher ability to conduct firm on-the-job search.[28] Unlike the case where improvements in λ_w also led to wages growing at a faster rate with improvements in x, a higher λ_f can lead to wages growing more slowly in x. Notice that even if wages are higher in levels from a higher λ_f, but so long as productivity is growing much faster than the increase in wages, a productivity-wage gap can emerge. Thus, even though there is more mass at higher productivity levels, a higher λ_f simultaneously causes wages to grow more slowly with improvements in x, encouraging the emergence of a productivity-wage gap.

$$w(x) = \eta x + (1-\eta)\widetilde{x} - (\rho+\delta)\chi + (1-\eta)\lambda_f q \int_x^{\overline{x}} S(y)\pi(y)\,dy + (1-\eta)^2 \lambda_f q \int_{\widetilde{x}}^{x} S(y)\pi(y)\,dy \tag{24}$$

[27] See equation 36 in Appendix B.4 for a formal proof

[28] See Appendix B.4 for a formal proof that the slope of the wage function is decreasing in λ_f.

$$w'(x) = \eta \left[1 - (1-\eta)\lambda_f q \pi(x) S(x)\right] \tag{25}$$

4.3 The Slope of the Wage Function

The previous comparative static exercises were informative in terms of helping us understand the role of firm vs worker on-the-job search in terms of affecting average labor productivity and average wages. Whether the productivity-wage gap diverges or not depends on how wages change with productivity. Equation 26 shows that when both firm and worker on-the-job search are operative, i.e. $\lambda_f, \lambda_w \neq 0$, the ability of workers to conduct on-the-job search relative to firms' ability affects the rate at which wages grow with improvements in x.

$$w'(x) = \eta + (1-\eta)\eta \left[\lambda_w p \frac{v^u}{v} - \lambda_f q \frac{u}{\ell}\right] S(x) \pi(x) \tag{26}$$

Because, in equilibrium both $\widetilde{x}$ and θ are changing with changes in λ_f and λ_w, we now turn to solving our model numerically to identify the extent to which firm and worker on-the-job search can affect the productivity-wage gap.

5 Numerical Exercise

5.1 Calibration

Our main goal is to examine if our model calibrated to match labor market flow rates can capture the increase in the productivity-wage gap. Since the productivity-wage gap only diverged post 2000, we conduct two separate exercises and split the time period into a pre-2000 period (1990-1999) and a post-2000 (2000-2017) period. We separately calibrate the parameters of our model to each time period's labor market flows.

A period in our model is a month. Accordingly, the discount rate, ρ, is set to 0.004 to reflect an annual interest rate of about five percent. Although a period in our model is a month, we time-aggregate our model-generated moments to a quarterly frequency. We do so since the information about the replacement hiring share is measured at quarterly frequency. The quarterly time aggregation allows us to capture replacement hires conducted for both the purposes of re-filling vacated positions as well as for firm on-the-job search reasons. Following the literature, we set the bargaining weight, η, to be 0.5. We use information from the Business Employment Dynamics (BED) database and compute the average monthly job destruction rate, δ for the period 1990-2017 to be 0.0235.[29] We hold constant the job destruction rate over the two time periods.

The empirical literature has typically found that wages are log-normally distributed. Thus, we assume that the distribution of match quality, $\Pi(x)$ is given by the log of $N(-\sigma_x^2/2, \sigma_x^2)$. Overall, this leaves us with six key parameters to calibrate $\{\lambda_f, \lambda_w, b, s, \sigma_x, \alpha\}$. We treat the fixed entry cost of job

[29] The job destruction rate from the BED is computed on a quarterly basis and is calculated as the sum of all jobs lost in either closing or contracting establishments divided by total employment. For the time period of interest, the average quarterly job-destruction probability is 0.068. To convert this number to a monthly rate, we calculate δ as

$$\delta = 1 - \exp(0.068/3)$$

creation χ as a residual to be solved for within the model. Further, we calibrate α, the elasticity of the meeting function with respect to unemployment to information on labor market flows from the first time period (pre 2000) and hold it constant at its calibrated value for the second time period.

While all key parameters are jointly calibrated, we use the following moments to identify the parameters: to pin down λ_f, we target the quarterly replacement hiring share in each period. To pin down λ_w, we target the quarterly employment-to-employment (EE) hiring share,[30] while we target an unemployment insurance ratio of 0.7 to pin down b, the value of home production. We use information on the quarterly rate at which employed individuals exit into unemployment (EU rate), and the unemployment rate to pin down s and σ_x. Finally, to pin down α, we use information on the quarterly job-finding rate (UE rate). Because there is a distinction between gross and net flows in our model, targeting the exit rate from employment, the job-finding rate and the unemployment rate does not automatically make any one of the moments a linear combination of the others. The exit rate from employment is a function of total gross separations which is composed of both exogenous separations s and endogenous separations stemming from firm on-the-job search. The job-finding rate out of unemployment is affected by the rate at which the unemployed worker meets and is hired by both unfilled and filled vacancies. In contrast, the unemployment rate is unaffected by the rate at which an unemployed worker is hired by a filled vacancy since the hiring of this unemployed applicant requires an employed worker to be displaced into unemployment, giving rise to no change in the unemployment pool on net.

Using data from the Current Population Survey (CPS) on employment, unemployment and short term unemployment, we find, for the period 1990m1-2000m1,[31] that the average monthly exit probability of an employed individual is about 0.032 while the average monthly job finding probability of an unemployed individual is given by 0.44.[32] In continuous time, this would imply that workers leave employment with at a quarterly rate of $-3*\log(1-0.032)$ and find jobs at a rate of $-3*\log(1-0.44)$. The average unemployment rate during this period is about 5.4%. For the pre-2000 period, we find that the quarterly EE hiring share is about 48%.[33] Using data from the QWI, we calculate that the average share of replacement hires for the period 1990Q1 to 1999Q4 is about 0.35.[34] For the period post-2000, we find, using CPS data, an average unemployment rate of 6.1%, an average monthly exit probability of 0.023, an average job-finding probability of 0.32 and an EE hiring share of about 0.41. From the QWI, we find that the replacement hiring share in the latter time period is about 0.38.

Table 1 summarizes our calibrated parameters. A few items are note-worthy from Table 1. Firstly and in line with the observed long-run trend decline in the hiring rate, our model requires that χ

[30] In calculating the quarterly EE hiring share, we include in the numerator all monthly EE hires as well as E-U-E hires that occur in a quarter.

[31] Our calibrated parameters would not change much if we expand the sample to be between 1951m1 to 1999m1 as the average unemployment rate during that period is about 5.5%, the exit probability is about 0.034 and the job finding probability is about 0.45. The targeted labor market moments for the period 1990m1-1999m1 is not substantially different from the period 1951m1 to 1999m1.

[32] We calculate the unemployment outflow and inflows rates by following the method proposed in Shimer (2012).

[33] We calculate the EE hiring share for the period 1994 onwards since that is when the CPS introduced a question that allows us to track whether the employed individual was in a same or different job. This number is consistent with information reported in Fallick and Fleischman (2004) who find that for the period 1994-2003, nearly 40% of all new hires observed an employer change, i.e. an EE transition.

[34] It is not inconsistent for the EE hiring share to be larger than the replacement hiring share. Note that not all EE transitions require a replacement hire to be conducted. A worker may quit its existing firm for a different job and the firm may not re-fill the position. In that case, an EE hire would be recorded but no replacement hire would be observed.

Table 1: Model Parameters

Fixed Parameters			
Parameter	Description		Value
ρ	discount rate		0.004
η	bargaining weight		0.5
δ	job destruction rate $1-\exp(-0.068/3)$		0.0235
Calibrated Parameters (1990-1999)			
Parameter	Value	Quarterly Targets	Model Moment
λ_f	0.181	replacement hiring share of 0.35	0.34
λ_w	0.087	EE hiring share of 0.48	0.46
b	0.611	70% UI ratio	0.68
s	0.019	exit rate of 0.098	0.097
σ_x	0.022	unemployment rate of 0.054	0.054
α	0.315	UE rate of 1.74	1.77
χ	3.690	residual from free-entry eqn	
Calibrated Parameters (post-2000)			
Parameter	Value	Quarterly Targets	Model Moment
λ_f	0.076	replacement hiring share of 0.38	0.38
λ_w	0.019	EE hiring share of 0.41	0.41
b	0.562	70% UI ratio	0.70
s	1.8e-8	exit rate of 0.070	0.070
σ_x	0.024	unemployment rate of 0.061	0.061
χ	7.514	residual from free-entry eqn	

increase across the two time periods so as to reduce entry and capture the more muted UE rates in the second time period. This is qualitatively in line with the decline in firm entry rates observed in the data. Our model predicts that the creation of new job positions would have declined by 11% acrossed the two time periods.[35] This lower entry is consistent with the decline in firm entry over this period of time. Establishment openings in the BED declined by about 26% across these two time periods while establishment entry rate as measured in the Business Dynamics Statistics (BDS) declined by 18%

Secondly, because replacement hiring and EE transitions as a percentage of total employment are also falling across the two time periods, our calibrated exercise also requires λ_f and λ_w to fall across time with the decline in hiring rates. Notably, the extent to which λ_w falls is much greater than λ_f. Post 2000, λ_w is about 22% its original value in the pre-2000 period while λ_f is about 42% its original value. This together with the rise in χ, which has the effect of reducing firm entry and raising q, suggests that relative to workers, firms were more able to conduct on-the-job search. Indeed, given these parameter values, we find that effective rate at which workers and firms could conduct on-the-job search in the pre-2000 period were given by $\lambda_w * p = 0.16$ and $\lambda_f * q = 0.13$ respectively. In contrast, in the post 2000 period, workers' and firms' effective rate at which they could conduct on-the-job search were given

[35] Using the laws of motion for unfilled vacancies and the unemployed, we can define new jobs as $v^{new} = \delta\,(v^u + 1 - u)$ and further define the rate at which new jobs are created as $v^{new}/(1-u)$.

by $\lambda_w * p = 0.02$ and $\lambda_f * q = 0.07$ respectively, suggesting that while both parties observed a decline in their ability to conduct on-the-job search, firm on-the-job search became more prevalent relative to worker on-the-job search in the second time period.

The declines in χ, λ_f and λ_w have implications for net employment growth - defined in the QWI as the difference between firm job gains and firm job losses. While we do not explicitly target the net employment growth in our model, our model predicts that net employment growth would have declined by 66%,[36] a value somewhat below the actual observed decline in net employment growth in the data of 95%.[37] Nonetheless, our calibrated values of λ_f, λ_w and χ across the two time periods largely capture the slowdown in hiring rates and net employment growth.

Thus far, our model has been calibrated to capture labor market flows in and out of unemployment as well as the replacement hiring share and EE hiring share. We now ask whether model calibrated to match labor market flows can capture the divergence in the productivity-wage gap post 2000.

5.1.1 Model Predictions

Having calibrated the model to match the observed rise in replacement hiring, we now examine how much the productivity-wage gap would have increased by. Table 2 shows our results. As aforementioned, all results in our model critically depend on equilibrium reservation match quality and labor market tightnes,$\{\widetilde{x}, \theta\}$. The increase in χ and declines in λ_f and λ_w causes $\widetilde{x}$ to increase across the two time periods. The higher selectivity over reservation match-quality as depicted by the rise in $\widetilde{x}$ suggests that there are more matched firm-worker pairs at higher levels of x. According to the model, the overall labor-productivity is higher in the post 2000 period with Y/N increasing by about 1 percentage point. Consistent with the lower UE rates we calibrate the model to match, labor market tightness, θ, falls across the two time periods by about 42%.

Coming to our main question of interest, we now explore our model's implications for the productivity-wage gap. We define the productivity-wage gap as the ratio of labor productivity, Y/N, to the average wage in the economy. In the data, the productivity-wage gap rose by about 8% between the two time periods. Our model predicts about a 12% increase across the two time periods. Across these two time periods, our model predicts that about one-twelfth of the increase in the productivity-wage gap comes from an increase in labor productivity, Y/N, and the rest of the widening of the gap is accounted for by the decline in average wages.

This widening in the productivity-wage gap largely stems from the fact that a higher ability of the firm to conduct on-the-job search, $\lambda_f q$, relative to the worker, $\lambda_w p$ affects the bargaining positions of both the firm and the worker. In particular, the higher ability of firms to conduct on-the-job search relative to workers raises the firm's outside option, χ, while simultaneously worsening the worker's outside option, U. To understand the decline in the worker's outside option, we look at two objects: 1) job insecurity, which we measure as the fraction of exits into unemployment that arise from endogenous

[36] In our model, job gains measure only hiring by unfilled vacancies since job gains is measured as $\mathbb{I}(\text{Hires} - \text{Replacement Hires}) > 0$. Job losses stem from both exogenous separations, s, and endogenous worker separations, $p^*(x)$, and from the job being destroyed, δ.

[37] Using the QWI data, net employment change is calculated as $\frac{\text{Job Gains} - \text{Job Loesses}}{\text{Employment}}$, where the measure of job gains and job losses is taken directly from the QWI.

Table 2: Non-targeted Model-Generated Moments

Pre vs. Post 2000				
	description	Pre-2000	Post 2000	Percent change
$\widetilde{x}$	reservation productivity	0.94	0.97	3.2
θ	labor market tightness	2.44	1.41	-42
$\lambda_f q$	firm OTJ contact rate	0.13	0.07	-49
$\lambda_w p$	worker OTJ contact rate	0.16	0.02	-85
job insecurity	fraction of EU that is endogenous	0.34	0.40	16
$\lambda_f v^m/v$	fraction of vacancies firm OTJ	0.50	0.64	28
ρU	worker outside option	0.88	0.79	-11
χ	firm outside option	3.59	7.53	110
Y/N	labor productivity	1.08	1.09	0.9
mean w	average wage	0.90	0.80	-11
$\frac{Y/N}{\text{mean} w}$	productivity-wage gap	1.19	1.35	12

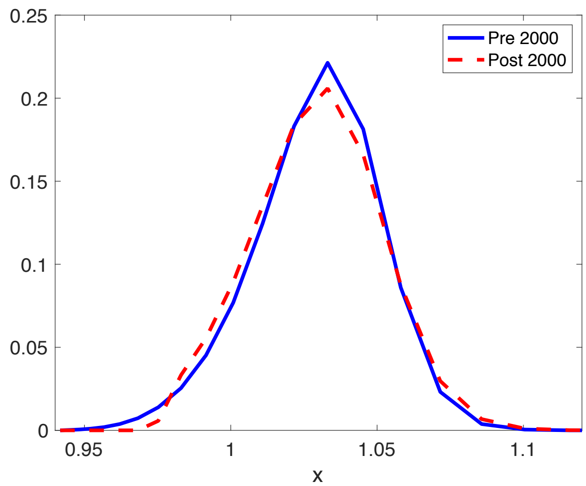
separations, i.e.

$$\text{job insecurity} \;=\; \frac{\int_{\widetilde{x}}^{\overline{x}} q^*(x)dF(x)}{\int_{\widetilde{x}}^{\overline{x}} q^*(x)dF(x) + s + \delta}$$

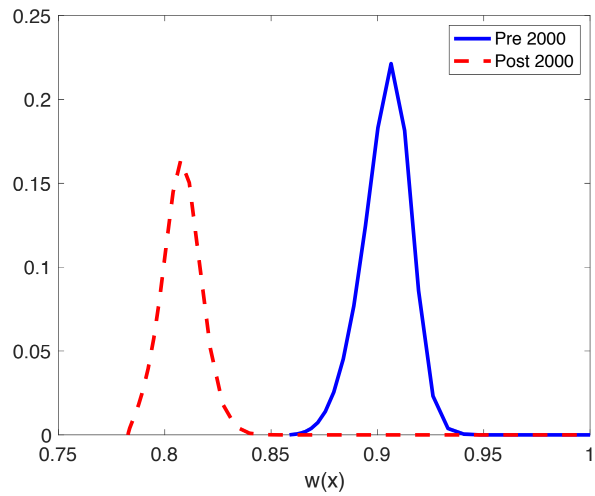
and 2) measured matching-efficiency which we proxy as the fraction of vacancies that are actually firms conducting on-the-job search, $\lambda_f v^m/v$.

Focusing first on job insecurity, we observe that precisely because firms have a higher relative rate of searching on-the-job than workers, the fraction of exits into unemployment for the worker that stem from endogenous separations rises in the post 2000 period rises by 16%. This higher job-insecurity lowers the value of employment for a worker as the worker is now more likely to be replaced. The lower employment value stemming from a shorter employment spell in turn reduces the worker's outside option and causes U to fall. At the same time, the higher rate of firm on-the-job search relative to worker on-the-job search also changes the composition of vacancies, tilting it in favor of vacancies made up of firms doing on-the-job search. In our model, the post 2000 period observes a 28% increase in the share of vacancies conducting firm on-the-job search, implying a decline in matching efficiency as applicants must draw a match-quality above the incumbent worker's match quality before they are employed. This decline in matching efficiency also contributes towards the worsening of workers' outside options.

This change in both the firm's and worker's outside options in turn affects how much of a higher x the firm passes on as wages to the worker. As foreshadowed in our comparative static exercises, a higher λ_w implies a larger $dw(x)/dx$ while a higher λ_f implies the opposite. When firms observe a higher rate of conducting on-the-job search relative to workers, our comparative static exercises suggest that wage growth may be slower in x. Figures 5a and 5b show how the productivity and wage distributions over the two time periods evolve. Focusing on the distribution of match-quality, Figure 5a shows that the productivity distribution in the post 2000 period has marginally more mass at higher productivities. This higher mass at higher match productivities stems from the fact that in equilibrium the declines in both λ_f and λ_w cause both firms and workers to be more selective (higher $\widetilde{x}$). Intuitively, when workers



(a) Productivity Distribution



(b) Wage Distribution

Figure 5: Model-implied Productivity and Wage Distribution

(firms) have little opportunity to re-match with another vacancy (applicant), they become more selective over their first match out of unemployment (out of being unfilled). As such, the declines in λ_f and λ_w imply that for a given market tightness, both worker and firm have a higher reservation match quality. This higher selectivity in turn leads to more productive matches being formed on average.

Although there was marginally more mass at higher productivities in the second time period, Figure 5b shows that wages were further skewed to the left despite the higher mass at the top. Again, this comes from the fact that although surplus is larger with higher x, workers' values are depressed through firms' higher relative ability to conduct on-the-job search.

Importantly, in our calibration exercise we did *not* target any wage moments in the data. Using CPS data, we compute the portion of wages not explained by observables and analyze how the skewness in residualized wages changed over time.[38] We find that in the data, the residual wage distribution became less negatively skewed, rising from -0.096 in the pre-2000 period to -0.002 in the post 2000 period. The reduction in negative skewness suggests a larger concentration at lower wages in the post 2000 period and is qualitatively consistent with the leftward shift in the wage distribution we observe here.[39] Taking 10000 random draws from our model's predicted wage distributions, we compute that the skewness of our wage distribution pre 2000 is about -0.26. This skewness turns positive post 2000 to a value of 0.19.[40] One might argue that the leftward shift in the wage distribution is seemingly counterfactual,

[38]Here, we run a mincer wage regression of log wages against education dummies, age and age-squared, a female dummy and race dummies, and compute the skewness of the residual wage distribution. We do this because our model has nothing to say about age gender, race and educational attainment, all of which are demographic factors that affect wages and that could changed over the two time periods.

[39]Our model predicts a leftward shift in the wage distribution whereas the wage regression by construction assumes that the residual term has an expected mean of zero. As such, we find it more useful to compare what our model implies in terms of skewness.

[40]Clearly because we did not target wage moments, and because the residualized wage from our regression has a mean of 0, the skewness values may not be the same between our model and the data. However, we emphasize that both our model and the data shares the feature that the wage distribution observed larger mass at lower wages across the two time periods.

since real wages are growing in the data albeit slowly. The correct way to interpret this shift down is relative to trend growth. However, it should be noted that if we had assumed a balanced growth path where the exogenous match quality distribution, $\Pi(x)$, were allowed to shift rightward over time, our model would also predict wages growing over time as opposed to declining. The key take-away of our model is the prediction that a slower growth rate of wages with productivity would result as firm on-the-job search becomes more prevalent relative to worker on-the-job search.

Having discussed our model's predictions for the productivity-wage gap, we turn now towards identifying the role of λ_w and λ_f in affecting the rise in the productivity-wage gap.

5.2 The role of worker vs. firm on-the-job search

A key question is how much of the rise in the productivity-wage gap can be attributed to the change in λ_w vs. the change in λ_f. To answer this question, we now conduct the following counterfactual exercises. We keep all other parameters fixed at their post 2000 values and ask how much larger the productivity-wage gap would be if either λ_w or λ_f were separately held at their pre-2000 values.[41]

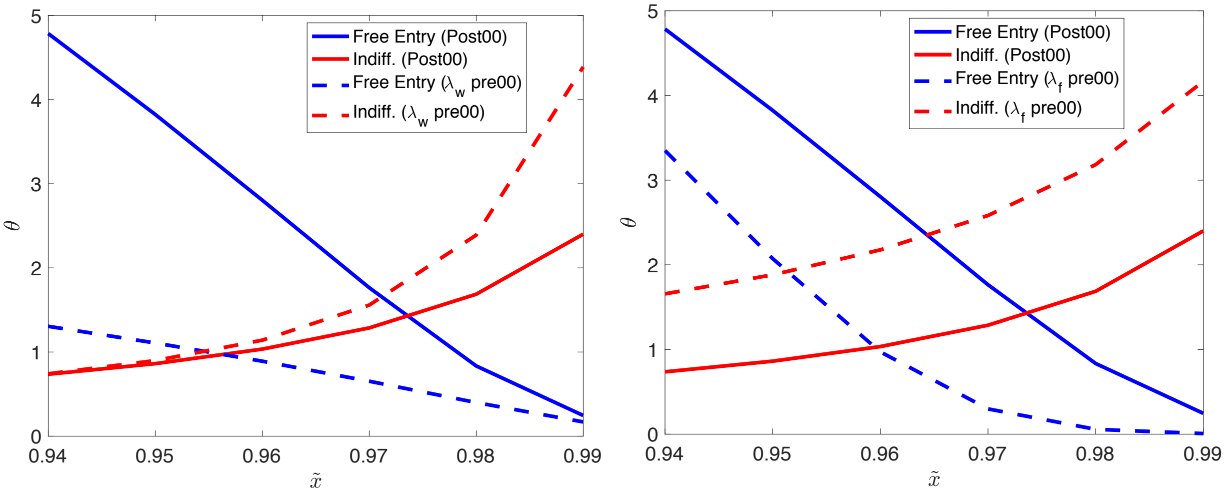
The role of worker on-the-job search: We begin first by holding λ_w at its pre-2000 value and allowing all other parameters to be updated to their post-2000 value. One argument for why the productivity-wage gap had widened is that worker on-the-job search had declined over time and as a result, the rate at which wages grow relative to productivity may be lower when workers are not able to climb the ladder as fast. This stems from the notion that a lower opportunity to climb the job ladder, i.e. lower λ_w, reduces both the value of being employed and unemployed, and this in turn reduces the wages that workers may receive.

Table 3: Counterfactual: λ_w fixed at pre 2000 level

λ_w counterfactual				
	description	λ_w pre-2000	Post 2000 (all)	Percent Diff
$\widetilde{x}$	reservation productivity	0.95	0.97	-1.9
θ	labor market tightness	1.00	1.41	-29
$\lambda_f q$	firm OTJ contact rate	0.08	0.07	11
$\lambda_w p$	worker OTJ contact rate	0.09	0.02	269
job insecurity	fraction of EU that is endogenous	0.36	0.40	-10
$\lambda_f v^m/v$	fraction of vacancies firm OTJ	0.51	0.64	-20
ρU	worker outside option	0.78	0.79	-0.5
Y/N	labor productivity	1.08	1.09	-0.3
mean w	average wage	0.80	0.80	-0.6
$\frac{Y/N}{\text{mean} w}$	productivity-wage gap	1.35	1.35	0.3

Table 3 shows our results. A few results are note-worthy. Firstly, keeping λ_w fixed at its higher pre-2000 value leads to a lower $\widetilde{x}$ and lower labor market tightness, θ. This is because when λ_w is high,

[41] In all our exercises, we keep χ fixed at its post 2000 value.



(a) λ_w pre 2000

(b) λ_f pre 2000

Figure 6: Equilibrium Under Our Counterfactual Exercises

firms are less able to retain workers as workers can now contact outside offers at a higher rate. This lower retention probability causes the free entry curve to shift leftward as depicted in Figure 6a. At the same time, because workers under a higher λ_w have increased opportunity to keep meeting vacancies while employed, workers' selectivity over reservation match quality declines, causing the worker indifference curve to shift upward. Because the shift in the free entry condition was larger than the change in the worker indifference curve, $\widetilde{x}$ and θ fall in equilibrium as depicted in Figure 6a. The lower reservation match quality in equilibrium results in a larger weight on lower productivity values, and as such lowers labor productivity by 0.3% relative to its post 2000 counterpart.

Next, we observe that under our counterfactual exercise, the rate at which workers can conduct on-the-job search, $\lambda_w p$, relative to firms $\lambda_f q$, is higher when λ_w is fixed at its pre-2000 value. This higher relative ability of workers to conduct on-the-job search in turn causes both job insecurity and the fraction of filled vacancies conducting firm on-the-job search to be lower than its post 2000 counterpart by 10 and 20 percent respectively.

Despite the decline in job insecurity and the rise in measured matching efficiency, workers' outside options still fall marginally by 0.5% when λ_w is fixed at its pre-2000 value. The reason for this is because labor market tightness also falls by 29% (relative to its post 2000 counterpart) under this counterfactual exercise, making it harder for unemployed and employed workers to contact vacancies on average. This decline in U as well as the lower $\widetilde{x}$ implies that the average wage remains roughly constant. Because both labor productivity and average wages do not vary much from their post 2000 counterpart, the productivity wage gap is largely unchanged.

Overall, we find that keeping λ_w at its higher pre-2000 value would have done little to affect the productivity-wage gap. On the one hand, the higher ability of workers to conduct on-the-job search relative to firms reduced job insecurity and measured matching efficiency which in turn would have raised the worker's outside option and the rate at which wages grow in improvements in x. On the other hand, the higher relative ability of workers to conduct on-the-job search caused both labor market tightness as well as firms' and workers' selectivity over reservation match quality to fall. The declines in $\widetilde{x}$ and θ counteract the positive gains from reduced job security and higher measured matching efficiency, causing workers' outside options to remain largely unchanged. This together with the marginally lower average labor productivity causes the productivity-wage gap to exhibit little to no difference. As such, we conclude from the results of our counterfactual exercise that the decline in worker's ability to do on-the-job search is not the key factor driving the divergence in the productivity-wage gap.

The role of firm on-the-job search: We repeat the same exercise as before but now hold λ_f to its pre-2000 value and update all other parameters to their post 2000 values. When λ_f is kept at its higher pre-2000 value, firms can easily replace workers, causing workers' bargaining position or threat point to be reduced. As such, Figure 6b shows that when λ_f is kept at its pre-2000 value, the workers' indifference curve shifts upward as workers become less selective over their reservation match quality. At the same time, the higher opportunity of re-matching with applicants while having a filled position causes firms to be less selective over their reservation match quality and for the free entry curve to shift leftward. Overall, the higher pre-2000 λ_f causes reservation match quality to fall and labor market tightness to increase as depicted in Figure 6b

Table 4: Counterfactual: λ_f fixed at pre 2000 level

λ_f counterfactual				
	description	λ_f pre-2000	Post 2000 (all)	Percent Diff
$\widetilde{x}$	reservation productivity	0.95	0.97	-2.2
θ	labor market tightness	1.92	1.41	36
$\lambda_f q$	firm OTJ contact rate	0.14	0.07	111
$\lambda_w p$	worker OTJ contact rate	0.03	0.02	24
job insecurity	fraction of EU that is endogenous	0.53	0.40	34
$\lambda_f v^m/v$	fraction of vacancies firm OTJ	0.82	0.64	28
ρU	worker outside option	0.78	0.79	-0.2
Y/N	labor productivity	1.11	1.09	2.2
mean w	average wage	0.80	0.80	-0.6
$\frac{Y/N}{\text{mean} w}$	productivity-wage gap	1.39	1.35	2.5

Fixing λ_f to its pre-2000 value causes the rate at which firms can conduct on-the-job search to be close to five times the rate at which workers can conduct on-the-job search. This in turn raises the job insecurity and the fraction of filled vacancies that conduct firm on-the-job search to increase by 34 and 28 percent respectively. Both the rise in job insecurity and the decline in measured matching efficiency put downward pressure on workers' outside options. Thus despite a higher labor market tightness, workers' outside option, U, falls by 0.2%.

All these forces act towards raising the productivity-wage gap. As discussed earlier in our comparative static exercises, the match productivity distribution under a higher λ_f, $F(x \mid \lambda_f^{high})$ first order stochastically dominates the match productivity distribution under a low λ_f, $F(x \mid \lambda_f^{low})$. Because λ_f is high in the pre-2000 period, and because its effective rate of conducting on-the-job search

$\lambda_f q$ is even higher than the case where all parameters were kept at their pre-2000 values[42], the average labor productivity in this counterfactual economy actually rises. Intuitively, when firms have a higher absolute effective rate of re-matching with better workers, labor productivity increases as the productivity distribution exhibits more mass at higher match quality values. At the same time, the reduction in unemployment values cause average wages to decline slightly, attenuating the widening of the productivity-wage gap. Overall, our results suggest that had λ_f not also fallen in the post 2000 period, the divergence in the productivity-wage gap would have been even higher. Given our counterfactual results, we conclude that the extent of firm on-the-job search plays an important role in affecting the magnitude of the productivity-wage gap.

6 Discussion

Worker on-the-job search Recent work by Fujita et al. (2019) argues that employment-to-employment (EE) transitions have not actually declined post 2000 but are actually an artefact of an increase in missing answers surrounding the question on whether the individual was still with the same employer as the previous interview month.[43] To address the possibility that employment-to-employment transitions did not decline, we repeat our calibration exercise but make the targeted EE share in the post 2000 period equal to that observed in the data pre-2000. Appendix C contains our results. Importantly, relative to the post 2000 period where the EE share was allowed to fall, we find that fixing the EE share implies a higher λ_f and λ_w, and a lower χ for the post 2000 period. Because λ_f, λ_w and χ are all changing in our model recalibrated to match a fixed EE share, we isolate the effects of λ_w and λ_f separately by repeating the counterfactual exercises as described in Section 5.1.1. Table 8 in Appendix C shows our results. We find the same results as before, keeping λ_w fixed at its pre-2000 level does little to change the productivity-wage gap while keeping λ_f fixed at its pre-2000 value would have raised the productivity-wage gap by 2%. As such, our model's main finding that the productivity-wage gap depends on firms' ability to conduct on-the-job search relative to workers continues to hold even when we re-calibrate the model to match a constant EE hiring share across time periods.

Wage Determination We assume that wages are determined via Nash Bargaining and that firms (workers) must leave their incumbent workers (firms) prior to bargaining with their new applicants (employers). The results in our paper would not qualitatively change if we had instead assumed a different wage determination protocol such as that of sequential auctions as in Postel-Vinay and Robin (2002). Note that in this case, matched workers (firms) leave their incumbent firm (worker) whenever the vacancy (applicant) contacted offers the worker (firm) a value larger than what its incumbent firm (worker) can offer to retain her. For matches where the vacancy contacted cannot counter the incumbent firm's highest offer, the offer from the vacancy contacted can still be used to bump up the worker's outside option and to renegotiate wages upward. In the same vein, when the firm does on-the-job search and contacts an applicant who cannot counter its incumbent worker's highest offer to retain

[42] In Table 2, the pre-2000 value of $\lambda_f q = 0.13$. In the counterfactual exercise where only λ_f is kept at its pre-2000 value, $\lambda_f q = 0.14$

[43] This does not mean that the hiring rate has not declined since using data from the QWI, we show that the hiring rate i.e. hires over employment was declining over time.

the firm, the potential match with the new applicant can be used to renegotiate the incumbent worker's wage down. As such, the prospect of firm on-the-job search still promotes a widening productivity-wage gap. In the sequential auctions framework, it comes through renegotiating incumbent workers' wages down, while in our Nash bargaining protocol it comes through changing workers' outside options.

Vacancy duration and the relevant measure of market tightness The standard DMP model assumes that vacancies expire instantaneously, implying that only the current flow of vacancies are relevant for computing market tightness. In contrast, we argue that our assumption of long-lived job positions better accords with how the data on job openings is collected. In particular, the Bureau of Labor Statistics (BLS) which conducts the monthly JOLTS states that the information it collects on job openings are "a stock,or point-in-time, measurement for the last business day of each month".[44] Further, our measure of vacancies is not inconsistent with the JOLTS definition of a job-opening. Specifically, JOLTS requires a job opening to satisfy three criteria: 1) a position exists, 2) work can start in 30 days, and 3) the firm is actively recruiting where active recruiting implies that the firm has undertaken "steps to fill a position". Firms who conduct on-the-job search in our model satisfy these three conditions. Our model calls to attention that the proportion of unfilled vacancies in addition to labor market tightness are important for rationalizing job-finding rates.

Additional hiring costs Our paper also assumes a fixed cost of creating job positions and abstracts from other types of hiring costs. Inclusion of a hiring cost paid at the time of matching with a worker would not qualitatively change the predictions of our model. Its inclusion would only serve to raise hiring thresholds. So long as the replacement hiring share increases, this will still drive a larger productivity-wage gap. Moreover, the one time fixed vacancy posting cost χ, rather than the standard flow cost of a vacancy, can be interpreted as the cost of maintaining a human resources department for the duration of the job position and therefore subsumes the additional costs that could have been incurred at the time of matching.

7 Conclusion

We document that the replacement hiring share in the US has risen over time alongside a widening productivity wage-gap. We develop a model that incorporates both worker and firm on-the-job search and examine the implications that worker vs. firm on-the-job search has for productivity and wages. We find that, holding all else constant, both firm and worker on-the-job search cause wages and productivity to increase as workers and firms climb the ladder for better matches. However, the rate at which wages grow with improvements in x depends critically on whether worker vs. firm on-the-job search if more prevalent. Notably, firm on-the-job search, by worsening workers' outside options, can reduce the slope of the wage function, encouraging productivity to potentially grow at a faster rate than wages and thus giving rise to a productivity-wage gap. Numerically, we show that if firm's ability to conduct on-the-job search stayed at its pre-2000 levels, the productivity-wage gap would have been even higher by about 2.5%.

[44] See the JOLTS chapter of the "BLS Handbook of Methods". https://www.bls.gov/opub/hom/pdf/homch18.pdf

References

Albrecht, James, Bruno Decreuse, and Susan Vroman, "Directed Search with Phantom Vacancies," AMSE Working Papers 1707, Aix-Marseille School of Economics, Marseille, France March 2017.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen, "The Fall of the Labor Share and the Rise of Superstar Firms," NBER Working Papers 23396, National Bureau of Economic Research, Inc May 2017.

Azar, Jose, Ioana Marinescu, and Marshall I. Steinbaum, "Labor Market Concentration," NBER Working Papers 24147, National Bureau of Economic Research, Inc December 2017.

Burgess, Simon, Julia Lane, and David Stevens, "Job Flows, Worker Flows, and Churning," *Journal of Labor Economics*, July 2000, *18* (3), 473–502.

Cheron, Arnaud and Bruno Decreuse, "Matching with Phantoms," *Review of Economic Studies*, 2017, *84* (3), 1041–1070.

Davis, Steven J. and Brenda Samaniego de la Parra, "Application Flows," Technical Report 2017.

Elsby, Michael, Bart Hobijn, and Aysegul Sahin, "The Decline of the U.S. Labor Share," *Brookings Papers on Economic Activity*, 2013, *44* (2 (Fall)), 1–63.

Elsby, Michael W. L., Ryan Michaels, and David Ratner, "Vacancy Chains," Technical Report 2016.

Fallick, Bruce C. and Charles A. Fleischman, "Employer-to-employer flows in the U.S. labor market: the complete picture of gross worker flows," Technical Report 2004.

Fujita, Shigeru and Garey Ramey, "Job matching and propagation," *Journal of Economic Dynamics and Control*, November 2007, *31* (11), 3671–3698.

_ , Giuseppe Moscarini, and Fabien Postel-Vinay, "Measuring Employer-to-Employer Reallocation," Technical Report 2019.

Haefke, Christian and Michael Reiter, "Long Live the Vacancy," Technical Report 2017.

Karabarbounis, Loukas and Brent Neiman, "The Global Decline of the Labor Share*," *The Quarterly Journal of Economics*, 2013, *129* (1), 61.

Kiyotaki, Nobuhiro and Ricardo Lagos, "A Model of Job and Worker Flows," *Journal of Political Economy*, October 2007, *115* (5), 770–819.

Menzio, Guido and Espen R. Moen, "Worker replacement," *Journal of Monetary Economics*, September 2010, *57* (6), 623–636.

Mercan, Yusuf and Benjamin Schoefer, "Jobs and Matches: Quits, Replacement Hiring, and Vacancy Chains," Technical Report May 2019.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, November 2002, *70* (6), 2295–2350.

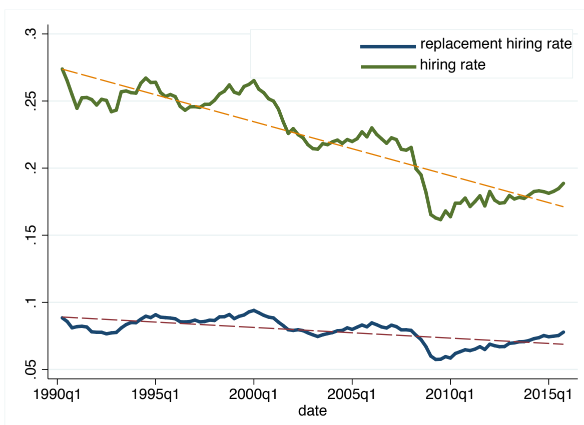
Shimer, Robert, "Reassessing the Ins and Outs of Unemployment," *Review of Economic Dynamics*, April 2012, *15* (2), 127–148.

Appendix

A Additional Data Work

A.1 Examining Replacement Hires

Figure 7a plots the replacement hiring rate against the hiring rate over time. The hiring rate is defined as the ratio of total hires to total employed while the replacement hiring rate is given by the ratio of replacement hires to total employed. Figure 7a highlights that higher replacement hiring share is a result of hires declining faster than replacement hires.



(a) Replacement vs. total hiring rate

A.2 Shift Share Analyses

To see if compositional changes in sectors are the primary factor behind the rise in the replacement hiring share, we conduct a shift-share analysis. In examining whether compositional changes were behind the rise in the replacement hiring share, we cut the data separately by firm age, firm size and by industry. Data by firm age and size are only reported for private firms in the QWI. We follow the firm age and size categories as provided by the QWI. For the shift share analysis done at the industry level, we use information available at the 2 digit NAICS level.

Because the divergence in the productivity-wage gap largely occurred after the 2000s, we divide the periods into pre-2000 and post 2000, and take the mean of the replacement hiring share in these two periods. We weight each firm age/size/industry category by their average employment share. Across the two time periods, the average replacement hiring share rose by about 3 percentage points. To assess how much of the increase in the replacement hiring share can be accounted for by compositional changes ("between") vs. by just an increase within each firm age/size/industry category, we use the following decomposition:

$$\Delta\text{Replacement Hiring Share} = \sum_{it} \Delta\left(rr_{it}\frac{emp_{it}}{emp_t}\right) \approx \underbrace{\sum_{it} \widehat{rr}\Delta\left(\frac{emp_{it}}{emp_t}\right)}_{\text{between}} + \underbrace{\sum_{it} \Delta rr_{it}\left(\widehat{\frac{emp_i}{emp}}\right)}_{\text{within}}$$

where Δ refers to the change, rr_{it} denotes the replacement hiring share of firm age/size/industry category i in period t, and emp_{it} denotes the employment in category i in period t. Terms with a "hat" denote averages. Table 5 shows the results from our shift share analysis:

Table 5: Shift-Share Analysis

Percent Explained by	Between	Within
By Firm Age	19.9%	80.1%
By Firm Size	26.7%	73.4%
By Industry	-0.0%	100%

While compositional changes at the firm age and firm size level explain close to a quarter of the changes in the replacement hiring share, a significant bulk of the increase in the replacement hiring share stems from changes within each firm age/size/industry category. This result is particularly stark at the industry level. Figure 8 shows that this is largely because the ranking of industries by employment share has not significantly changed, with the top 5 industries in the pre 2000 and post 2000 being exactly the same. In fact, our shift share analysis actually suggests that the change in industry composition actually contributes towards reducing the replacement hiring share by 0.005 points as opposed to increasing it.

Across firm age and size, we find that large and older firms were more likely to conduct replacement hiring. This is consistent with the notion that once firms reach their optimal firm size, they conduct replacement hires either for the purposes of finding a better worker or for the purposes of re-filling a position. Figures 9a and 9b show how replacement hiring shares vary across firm age and size for the two time periods.

B Surplus and Wages

Recruiting firm-worker pair Subtract (5) from (2) and rearrange to get:

$$(\rho + \delta + s + p^*(x))[J(x) - \chi] = x - w(x) + R(x) - (\rho + \delta)\chi \tag{27}$$

Next, subtracting ρU from both sides of (9) yields:

$$(\rho + \delta + s + q^*(x))[W(x) - U] = w(x) + H(x) - \rho U \tag{28}$$

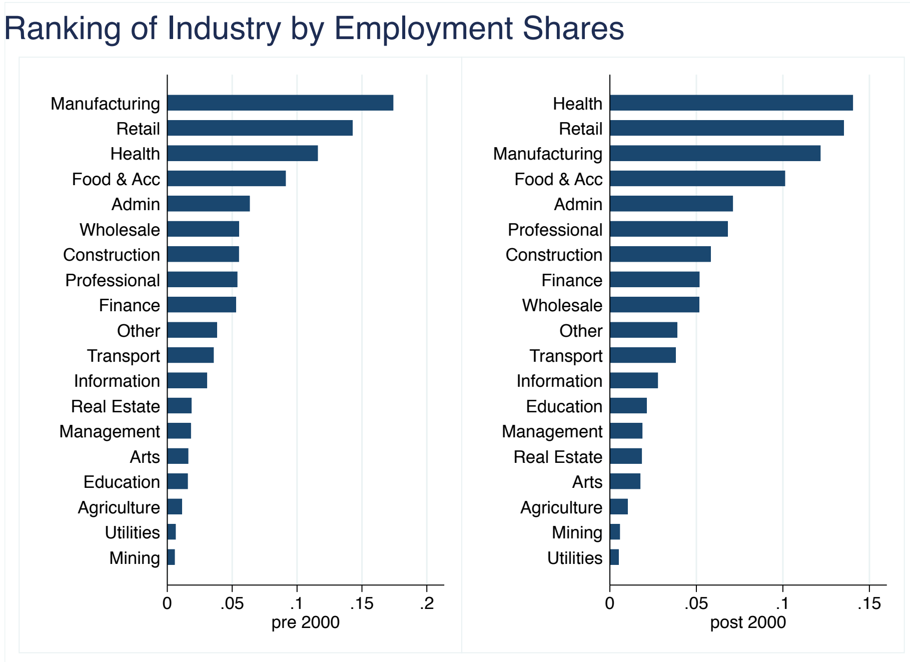


Figure 8: Ranking of industries roughly the same

Adding the two expressions and re-arranging gives us surplus:

$$\begin{aligned} \varrho(x) S(x) &= x - \rho U - (\rho+\delta)\chi \\ &+ \left[q(1-\eta)\frac{\lambda_f u}{\ell} + p\eta \frac{\lambda_w v^u}{v} \right] \int_x^{\overline{x}} S(y)\, d\Pi(y) \\ &+ q\frac{\lambda_f \lambda_w v^m}{\ell} \left\{ \int_x^{\overline{x}} S(y)\, d\Pi(y) F(x) + \int_x^{\overline{x}} \int_\epsilon^{\overline{x}} S(y)\, d\Pi(y)\, dF(\epsilon) \right\} \end{aligned} \tag{29}$$

where $S(x) = J(x) - J^u + W(x) - U$ and $\varrho(x) = \rho + \delta + s + p^*(x) + q^*(x)$. Next using the Nash Bargaining solution (14), the free entry condition (i.e. equation 5) can be re-written as:

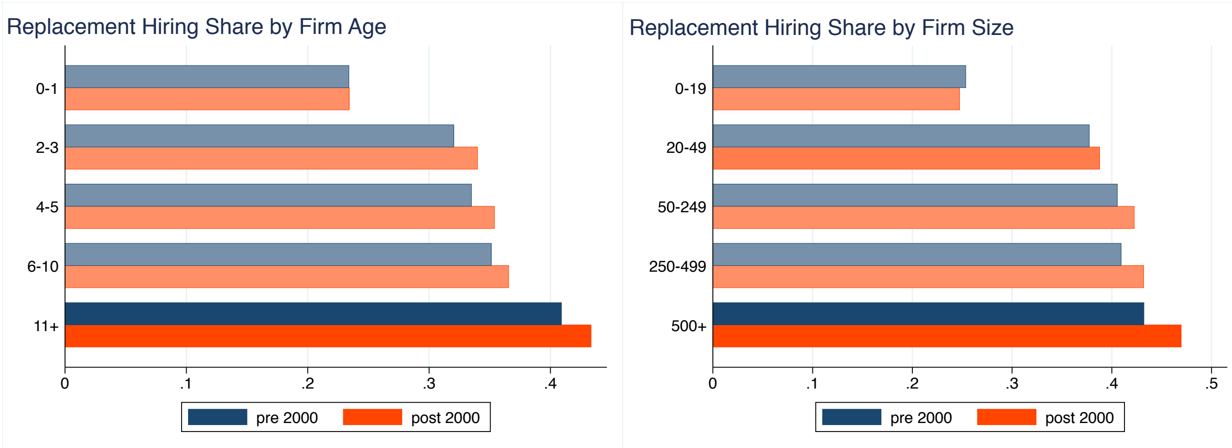
$$(\rho+\delta)\chi = (1-\eta) q \left[\frac{u}{\ell} \int_{\widetilde{x}}^{\overline{x}} S(y)\, d\Pi(y) + \lambda_w \frac{1-u}{\ell} \int_{\widetilde{x}}^{\overline{x}} \int_\epsilon^{\overline{x}} S(y)\, d\Pi(y)\, dF(\epsilon) \right]$$

Evaluating 29 at $\widetilde{x}$ and plugging for free-entry, we get equation 19.

B.1 General Wage Form

Note from the worker's gain to matching, we have:

$$(\rho + \delta + s + q^*(x)) \eta S(x) = w(x) + H(x) - \rho U$$



(a) By Firm Age

(b) By Firm Size

Figure 9: Replacement Hiring Shares by Firm Age and Size

Using the form of $H(x)$ and making $w(x)$ the subject of the equation, we have:

$$\begin{aligned} w\left(x\right) &= \eta x + \left(1-\eta\right)\rho U - \eta\left(\rho+\delta\right)\chi \\ &\quad + \left(1-\eta\right)\eta\left[\lambda_f q\frac{u}{\ell} - \lambda_w p\frac{v^u}{v}\right]\int_x^{\overline{x}} S\left(y\right)d\Pi\left(y\right) \end{aligned} \tag{30}$$

B.2 Distribution of productivity

Focusing on the measure of matched firm-worker pairs with match quality less than or equals to x and dividing everywhere by $v^m = (1-u)$, we get:

$$\begin{aligned} \left[q\frac{v^u}{(1-u)}\frac{u}{\ell}\right]\left[\Pi\left(x\right)-\Pi\left(\widetilde{x}\right)\right] &= \left(s+\delta\right)F\left(x\right) + qF\left(x\right)\left[1-\Pi\left(x\right)\right]\left[\lambda_f\frac{u}{\ell}+\lambda_w\frac{v^u}{\ell}\right] \\ &\quad +2\lambda_f qF\left(x\right)\frac{\lambda_w F\left(x\right)\left(1-u\right)}{\ell}\left[1-\Pi\left(x\right)\right] \\ &\quad +2\lambda_f qF\left(x\right)\frac{\lambda_w\left(1-u\right)}{\ell}\int_x^{\overline{x}}\left[1-\Pi\left(\epsilon\right)\right]f\left(\epsilon\right)d\epsilon \\ &\quad +\lambda_f q\frac{\lambda_w\left(1-u\right)}{\ell}\int_{\widetilde{x}}^{x}\int_z^{x}\left[\Pi\left(x\right)-\Pi\left(\epsilon\right)\right]f\left(\epsilon\right)f\left(z\right)d\epsilon dz \\ &\quad +\lambda_f q\frac{\lambda_w\left(1-u\right)}{\ell}\int_{\widetilde{x}}^{x}\left[\Pi\left(x\right)-\Pi\left(z\right)\right]F\left(z\right)f\left(z\right)dz \end{aligned} \tag{31}$$

B.3 Comparative Static: Worker On-the-job Search Only

Setting $\lambda_f = 0$, and using equation 31, we can show that the cumulative distribution of firm-worker pairs with match quality less than or equals to x becomes:

$$F(x) = \left[q\frac{v^u}{(1-u)}\frac{u}{\ell}\right]\frac{\left[\Pi\left(x\right)-\Pi\left(\widetilde{x}\right)\right]}{\left(s+\delta+p\lambda_w\left[1-\Pi\left(x\right)\right]\right)}$$

From the law of motion for the unemployed, we have:

$$q\frac{v^u}{1-u}\frac{u}{\ell}=\frac{s+\delta}{\left[1-\Pi\left(\widetilde{x}\right)\right]}$$

plugging this into $F(x)$, we get:

$$F\left(x\right)=\frac{s+\delta}{\left(s+\delta+p\lambda_w\left[1-\Pi\left(x\right)\right]\right)}\frac{\Pi\left(x\right)-\Pi\left(\widetilde{x}\right)}{1-\Pi\left(\widetilde{x}\right)}$$

From equation 30 when $\lambda_f=0$, the wage of type x also becomes

$$w\left(x\right)=\eta x+\left(1-\eta\right)\rho U-\left(1-\eta\right)\eta\lambda_w p\int_x^{\overline{x}}S\left(y\right)d\Pi\left(y\right)-\eta\left(\rho+\delta\right)\chi$$

and ρU is now given by:

$$\rho U=\widetilde{x}-\left(\rho+\delta\right)\chi+\eta\lambda_w p\int_{\widetilde{x}}^{\overline{x}}S\left(y\right)\pi\left(y\right)dy$$

Plugging in for ρU expression in $w(x)$, we get:

$$w\left(x\right)=\eta x+\left(1-\eta\right)\widetilde{x}-\left(\rho+\delta\right)\chi+\left(1-\eta\right)\eta\lambda_w p\int_{\widetilde{x}}^{x}S\left(y\right)\pi\left(y\right)dy \tag{32}$$

At this point it is useful to note that from equation 29 becomes:

$$\left(\rho+s+\delta+\lambda_w p\left[1-\Pi\left(x\right)\right]\right)S\left(x\right)=x-\rho U+\eta\lambda_w p\int_x^{\overline{x}}S\left(y\right)d\Pi\left(y\right)-\left(\rho+\delta\right)\chi$$

Differentiating the above by x and re-arranging, we get:

$$-\frac{\left(1-\eta\right)\lambda_w p\pi\left(x\right)S\left(x\right)}{\rho+s+\delta+\lambda_w p\left[1-\Pi\left(x\right)\right]}+\frac{dS\left(x\right)}{dx}=\frac{1}{\rho+s+\delta+\lambda_w p\left[1-\Pi\left(x\right)\right]}$$

The above is now an ordinary differential equation in x. So solving, we get:

$$S\left(x\right)=\frac{1}{\left(\rho+s+\delta+\lambda_w p\left[1-\Pi\left(x\right)\right]\right)^{1-\eta}}\int_{\widetilde{x}}^{x}\frac{1}{\left(\rho+s+\delta+\lambda_w p\left[1-\Pi\left(y\right)\right]\right)^{\eta}}dy \tag{33}$$

Plug 33 into 32 and we arrive at 34:

$$w(x)=\eta x+\left(1-\eta\right)\widetilde{x}-\left(\rho+\delta\right)\chi+\left(1-\eta\right)\eta\int_{\widetilde{x}}^{x}\int_{\widetilde{x}}^{y}\frac{1}{\left(\frac{\rho+\delta+s}{\lambda_w p}+1-\Pi\left(y\right)\right)^{1-\eta}\left(\frac{\rho+\delta+s}{\lambda_w p}+1-\Pi\left(z\right)\right)^{\eta}}dzd\Pi\left(y\right) \tag{34}$$

Differentiating the above wrt x, we get the slope of our wage function:

$$w'\left(x\right)=\eta+\left(1-\eta\right)\eta\pi\left(x\right)\left[\int_{\widetilde{x}}^{x}\frac{1}{\left(\frac{\rho+\delta+s}{\lambda_w p}+\left[1-\Pi\left(x\right)\right]\right)^{1-\eta}}\frac{1}{\left(\frac{\rho+\delta+s}{\lambda_w p}+\left[1-\Pi\left(y\right)\right]\right)^{\eta}}dy\right]$$

where one can show that the above slope of the wage function is strictly increasing in λ_w.

B.4 Comparative Static: Firm On-the-job Search Only

Shutting down worker on-the-job search, $\lambda_w = 0$, one can show using equation 31 and the law of motion for the unemployed, that the productivity distribution for matched firm-worker pairs with match quality $\leq x$ is given by:

$$F(x) = \left(\frac{s+\delta}{s+\delta+\lambda_f q\left[1-\Pi(x)\right]}\right)\frac{\Pi(x)-\Pi(\widetilde{x})}{1-\Pi(\widetilde{x})}$$

Turning to wages and using equation 30, under $\lambda_w = 0$, we get

$$w(x) = \eta x + (1-\eta)\rho U + (1-\eta)\eta\lambda_f q\frac{u}{\ell}\int_x^{\overline{x}} S(y)\, d\Pi(y) - \eta(\rho+\delta)\chi$$

Note that in this case, we have

$$\rho U = \widetilde{x} - (\rho+\delta)\chi + q(1-\eta)\lambda_f\int_{\widetilde{x}}^{\overline{x}} S(y)\pi(y)\, dy$$

Plugging in for ρU expression in $w(x)$, we get equation 24, replicated below for ease of reference:

$$\begin{aligned} w(x) &= \eta x + (1-\eta)\widetilde{x} - (\rho+\delta)\chi \\ &\quad + (1-\eta)\lambda_f q\int_x^{\overline{x}} S(y)\pi(y)\, dy \\ &\quad + (1-\eta)^2\lambda_f q\int_{\widetilde{x}}^{x} S(y)\pi(y)\, dy \end{aligned}$$

At this point it is useful to note that from equation 29 becomes:

$$\begin{aligned} \left(\rho+\delta+s+\lambda_f q\left[1-\Pi(x)\right]\right)S(x) &= x - \rho U - (\rho+\delta)\chi \\ &\quad + (1-\eta)\lambda_f q\int_x^{\overline{x}} S(y)d\Pi(y) \end{aligned}$$

Again we can differentiate the above wrt x and solve the ordinary differential equation. In this case, we would arrive at:

$$S(x) = \int_{\widetilde{x}}^{x}\frac{1}{\left(\rho+s+\delta+\lambda_f q\left[1-\Pi(x)\right]\right)^{\eta}\left(\rho+s+\delta+\lambda_f q\left[1-\Pi(y)\right]\right)^{1-\eta}}dy \tag{35}$$

plugging in 35 into , we arrive at equation 36:

$$\begin{aligned} w(x) &= \eta x + (1-\eta)\widetilde{x} - (\rho+\delta)\chi \\ &\quad + (1-\eta)\int_x^{\overline{x}}\int_{\widetilde{x}}^{y}\frac{1}{\left(\frac{\rho+s+\delta}{\lambda_f q}+\left[1-\Pi(y)\right]\right)^{\eta}\left(\frac{\rho+s+\delta}{\lambda_f q}+\left[1-\Pi(z)\right]\right)^{1-\eta}}dz\pi(y)\, dy \\ &\quad + (1-\eta)^2\int_{\widetilde{x}}^{x}\int_{\widetilde{x}}^{y}\frac{1}{\left(\frac{\rho+s+\delta}{\lambda_f q}+\left[1-\Pi(y)\right]\right)^{\eta}\left(\frac{\rho+s+\delta}{\lambda_f q}+\left[1-\Pi(z)\right]\right)^{1-\eta}}dz\pi(y)\, dy \end{aligned} \tag{36}$$

An inspection of 36 makes clear that the wage $w(x)$ is increasing in λ_f.

Differentiating the equation 36, we get:

$$w'(x) = \eta - \int_{\widetilde{x}}^{x} \frac{\eta(1-\eta)\pi(x)}{\left(\frac{\rho+s+\delta}{\lambda_f q} + 1 - \Pi(x)\right)^{\eta} \left(\frac{\rho+s+\delta}{\lambda_f q} + 1 - \Pi(y)\right)^{1-\eta}} dy$$

which demonstrates that the slope of the wage function is decreasing in λ_f.

C Recalibrated Model: Fixed EE Share

C.1 Robustness Check: Fixed EE share

Following Fujita et al. (2019) who argue that EE rates have not actually declined over time, we fix our targeted EE share in the post 2000 period to be equal to the EE share in the pre-2000 period. Table 6 shows our calibration results from this exercise. Notably, fixing the EE share elevates both the level of λ_f and λ_w relative to what we observed in our benchmark post 2000 calibration, while χ is actually lower.

Table 6: Model Parameters, Fixed EE share

Calibrated Parameters (post-2000,Fixed EE share)			
Parameter	Value	Quarterly Targets	Model Moment
λ_f	0.097	replacement hiring share of 0.38	0.36
λ_w	0.057	EE hiring share of 0.48	0.46
b	0.574	70% UI ratio	0.70
s	2.4e-8	exit rate of 0.070	0.069
σ_x	0.027	unemployment rate of 0.061	0.061
χ	7.009	residual from free-entry eqn	

Further, Table 7 shows how large the productivity-wage gap would be if we had instead calibrated the model to target a constant EE hiring share across the two time periods. While the productivity-wage gap is narrower by 1.4% in this re-calibrated version of the model, it should be noted that all parameters changed under this re-calibration, making it hard to pinpoint the contribution of worker on-the-job search.

As such, Table 8 repeats the counterfactual exercises of Section 5.1.1 and demonstrates the effect of allowing workers vs. firms to have a higher opportunity of conducting on-the-job search. As before we find that if λ_f was higher, the productivity wage gap would have been larger whereas little changed under a higher λ_w.

Table 7: Counterfactual: EE share fixed at pre 2000 level

Fixed EE share counterfactual				
	description	EE share fixed	Post 2000	Percent Diff
$\widetilde{x}$	reservation productivity	0.96	0.97	-1.1
θ	labor market tightness	1.39	1.41	-0.9
$\lambda_f q$	firm OTJ contact rate	0.09	0.07	31
$\lambda_w p$	worker OTJ contact rate	0.07	0.02	215
job insecurity	fraction of EU that is endogenous	0.41	0.40	2.0
$\lambda_f v^m/v$	fraction of vacancies firm OTJ	0.61	0.64	-5.8
ρU	worker outside option	0.80	0.79	2.0
Y/N	labor productivity	1.09	1.09	0.6
mean w	average wage	0.82	0.80	2.0
$\frac{Y/N}{\text{mean} w}$	productivity-wage gap	1.33	1.35	-1.4

Table 8: Counterfactual: EE share fixed at pre 2000 level

Role of λ_w vs. λ_f (Fixed EE share)				
	description	λ_w	λ_f	Post 2000
$\widetilde{x}$	reservation productivity	0.96	0.95	0.96
θ	labor market tightness	1.16	1.63	1.39
$\lambda_f q$	firm OTJ contact rate	0.09	0.15	0.09
$\lambda_w p$	worker OTJ contact rate	0.09	0.08	0.07
job insecurity	fraction of EU that is endogenous	0.39	0.51	0.41
$\lambda_f v^m/v$	fraction of vacancies firm OTJ	0.56	0.75	0.61
ρU	worker outside option	0.80	0.80	0.80
Y/N	labor productivity	1.09	1.12	1.09
mean w	average wage	0.82	0.82	0.82
$\frac{Y/N}{\text{mean} w}$	productivity-wage gap	1.33	1.36	1.33